非参数统计

——方法与应用

易丹辉 编著



中国统计出版社

非参数统计

——方法与应用

易丹辉 编著

中国统计出版社

（京）新登字 041 号

图书在版编目（CIP）数据

非参数统计：方法与应用/易丹辉编著.
—北京：中国统计出版社，1995.6
ISBN 7-5037-1968-0

Ⅰ. 非…
Ⅱ. 易…
Ⅲ. ①非参数统计—方法 ②非参数统计—应用
Ⅳ. O212.7

中国统计出版社出版
（北京三里河月坛南街 38 号 100826）
新华书店经销
中国科学院印刷厂
850×1168毫米 32开本 9.625印张 2[illegible].5万字
1996年3月第1版 1996年3月北京第1次印刷
印数 1—3[illegible]册
定价：14.[illegible]元

前　　言

非参数统计方法是本世纪30年代中后期开始形成并逐步发展起来的。它是与"参数统计"相比较而存在，不依赖于总体分布及其参数，亦即不受分布约束的统计方法。作者在承担国家教委人文社会科学研究八五规划项目博士点基金项目"90年代我国居民消费结构及倾向的研究"过程中，感到在很多情况下，参数统计方法的运用受到限制，如研究居民消费行为、居民收入等级与消费结构的关系等等，从而转向了非参数统计方法的探讨。《非参数统计——方法与应用》作为该课题的一个成果奉献给读者，以求这一统计方法能在我国市场研究、社会研究以及医学、企业管理等领域研究及应用方面更加普及。

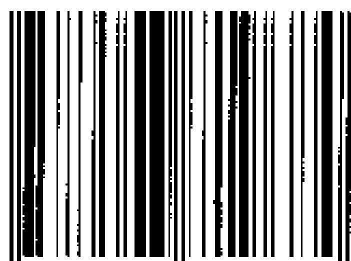
由于时间仓促、水平有限，书中不免有许多缺憾，恳请读者加以指正。

本书出版得到严建辉的支持，技术设计得到黄强、晏利东、范辉的大力帮助，作者在此一并表示感谢。

1994年11月30日

责任编辑： 吕 军
封面设计： 崔葆进

ISBN 7-5037-1968-0



定价：14.80 元

目　录

第一章　导　　言

一、测量的层次

所谓测量，就是根据一定的法则，给事物或事件分配一定的数字或符号。如测量职工对所从事职业的满意程度，可以分配0至10的数字，不满意给0分，很满意给10分，介于二者之间的分配中间的数字。于是，0至10这些抽象的数字符号就依据上述规定的法则，表示职工对所从事职业的满意程度。测量的作用在于准确地描述事物的类型、性质、状态，同时对事物之间的差异进行准确度量和比较。事物只有通过测量，才有可能选择适当的统计模型或公式，进行一系列的统计分析。

目前，广泛采用的测量层次为四种类型的尺度：定类尺度、定序尺度、定距尺度和定比尺度。

（一）定类尺度

定类尺度是按照事物的某些特征辨别和划分它们异同的一种测量层次，也被称作类别尺度、名义尺度。如性别、职业、民族等，都是按照事物的性质、类别区分的，均属定类尺度。

测量时，无论哪一种测量层次，都必须具有完备性和互斥性。完备性，是指用这种尺度测量某事物时，必须对这一事物所包括的各种情况都能进行测量。如测量性别时，应包括性别的各种情况：男、女，不能有遗漏。互斥性，是指用这种尺度测量时，不能有任一被测量对象跨越类别，即事物的各种情况具有互相排斥的不同值。如测量性别时，分为互斥的两类：男、女，某人或属于男，或属于女，不能既属于男又属于女。测量层次的完备性和互斥性，可以保证测量的准确无误。

定类尺度只能将事物分类，不能用以反映事物的数量状况，有

时，为了识别不同的类别，也用一定的数字和符号表示某类事物。如职工对所从事的职业是否满意，可以用"0"表示不满意。用"1"表示满意，这仅仅是人们赋予的识别标志，并不说明事物的数量。定类尺度是最低一个层次的测量尺度，它不能进行算术运算，而只能进行"="或"≠"的逻辑运算。定类测量数据的描述性统计量有：众数、频数等。

（二）定序尺度

定序尺度是按照事物的某种特征依顺序和级别进行排列的一种测量层次，也称作顺序尺度、等级尺度。例如，测量职工的文化程度可以采用定序尺度，分为：大专以上、中专或高中、初中、小学，显然前面的类别要比后面的类别高，即前面的文化程度最高，依次降低。定序尺度不仅能够区分事物，即对事物进行分类，而且可以反映事物在高低、大小、强弱上的差异，也就是使类别之间具有次序比较关系。定序尺度是比定类尺度高一层次的测量，它不仅能进行"="或"≠"的运算，还能进行">"、"<"的运算。最适合描述定序尺度中数据集中趋势的统计量是中位数，反映离散程度的是分位数。

（三）定距尺度

定距尺度是不仅能将事物区分类别和等级，而且可以确定其之间的数量差别、间隔距离的一种测量层次，也称作间隔尺度、区间尺度。例如，对学生学习成绩的测量，甲为 90 分，乙为 85 分，甲乙学生成绩间距为 90 − 85 = 5 分，这一测量就是定距尺度。定距尺度没有绝对的零点，也就是说，在这种测量中，任何两个间隔的差异与零点无关。例如，某门课程成绩的百分制测量，0 分不表示某考生没有这门课方面的知识，90 − 85 = 5 分，只表明甲生比乙生在这门课考试成绩中多 5 分，并不能说明甲生掌握的这方面知识是乙生的大约 1.06 倍(90/85)。定距尺度在实际应用中较为普遍，象温度、智商等都是定距测量。定距尺度是一种定量的测量层次，它不仅能反映事物的类别和顺序，而且能反映事物的具体数量

和数量之间的距离。它是比定序尺度又高一层次的测量，不仅能进行“＝”、“≠”，“＞”、“＜”的运算，还能进行“＋”、“－”的运算。定距尺度中描述性统计量，除了反映集中趋势的众数、中位数、均值外，还有反映离散程度的方差、标准差等，一般的定量统计方法都可以在这一测量层次应用。

（四）定比尺度

定比尺度是在定距尺度上增加绝对零点的一种测量层次，也称作等比尺度、比率尺度。例如对职工年龄的测量，这里 0 岁是非任意的，一个人年龄不可能比 0 岁更小，这一测量尺度对所有人都一样。若甲为 40 岁，乙为 20 岁，则甲的年龄是乙的 2 倍，这就是定比尺度的测量。是否具有实际意义的零点存在，是定比尺度与定距尺度的唯一区别。定比尺度由于有一绝对零点存在，因而比定距尺度更利于反映事物之间的比例或比率关系，它是所有测量层次中最高一层的测量，不仅能进行“＝”、“≠”，“＞”、“＜”，“＋”、“－”的运算，而且能进行“×”、“÷”的运算。在定比测量中，描述性统计量不仅有算术平均的均值，还有几何平均的均值，不仅有方差、均方差，还有变异系数等。

（五）四种测量尺度的关系

四种测量尺度有着不同的特点，其主要表现在作用和运算性质上，但它们之间又有较为密切的关系。首先，这些测量尺度之间有着包含关系，即高一层次的测量尺度总是包含低层次的测量尺度。定序尺度包含了定类尺度所有运算性质，定距尺度包含了定序、定类尺度所有运算性质，而定比尺度则包含了所有测量层次的运算性质。其次，四种测量尺度之间，低级的测量尺度往往能用较高级的测量尺度形式表示。例如，对学生考试成绩的测量，进行定类测量可分为及格、不及格；若将及格的成绩高低排序，可分为优、良、中、及格，这是定序尺度；若再将各顺序级给出等级分，则按百分制测量，优：90 分以上，良：80—90 分，中：70—80 分，及格：60—70 分。对同一事物的测量可以用多种尺度时，为避免信息的

丢失，应尽量将低层次测量尺度变成较高层次的尺度来测量。由于不同的测量层次具有不同的数学性质，因而在统计资料的收集、整理、分析过程中，往往需要采用不同的统计方法，也就是说，统计方法的运用总是与所选择的测量尺度相联系。在实际应用时，要首先弄清统计方法适用的测量尺度与所获得的资料采用的测量尺度是否一致。

二、统计检验

近代统计学的中心课题是统计推断。在统计推断中涉及这样的问题：如何利用部分事件的观察作出大量事件的结论。例如，要确定几种牌号的彩色电视机在我国居民中哪种最受欢迎，可以这样去搜集资料：到一家最大的商场站在柜台边，计数一天中每种牌号彩电的销售数量，几乎可以肯定那几种牌号彩电销量不同。但能否推断：那一天在这家商场销量最多的彩电是最受我国居民欢迎的呢？这取决于那种彩电的销售地域，也取决于那家商场的代表性，还取决于所观察的那些买主的代表性。统计检验正是要解决这一问题：如何根据样本值判断所得出的结论是否正确。

统计检验的一般步骤为：

陈述零假设(H_0)；

选择一种统计检验来检验 H_0；

给定显著性水平 α 和样本容量 n；

求出 H_0 成立时统计量的抽样分布；

确定否定域；

利用样本资料计算统计检验值。若其值落在否定域，则拒绝 H_0；若落在否定域之外，则在所选择的显著性水平上，不能拒绝 H_0。

(一) 零假设

零假设是一种无差别假设，表示要被拒绝的目的，也称原假设。备择假设(H_1)亦称对立假设，是与 H_0 相反的结论。若 H_0 被拒

绝，H_1 就可能被接受。例如，研究两种药物对治疗同一种病的效果不同。这个结论是要研究的假设，为了检验它，一般把它陈述为备择假设，可以用运算的形式写出为，$H_1: \mu_1 \neq \mu_2$，H_0 则是：$\mu_1 = \mu_2$，表明两种药物对治疗同一种病的效果相同。如果搜集的资料允许拒绝 H_0，则 H_1 被接受，这就支持了研究假设以及由它导出的理论。

H_1 的叙述是由研究假设的性质确定的。若研究假设只是考察两个事物有差异，则备择假设 H_1 为 $\mu_1 \neq \mu_2$；若考察其差值的方向，则 H_1 或者为 $\mu_1 > \mu_2$，或者为 $\mu_1 < \mu_2$。第一种检验为双尾检验（双侧检验），后两种检验为单尾检验（单侧检验）。

（二）显著性水平

假设检验是根据人们一条普遍的经验作为原则的，即小概率事件在一次实验中很难发生。如果一旦发生，就认为原来的假设不成立，从而拒绝 H_0。但是，很难发生并不等于决不发生，因此，在得出对 H_0 的判定时，可能会发生两类错误：第一类错误是当 H_0 实际上为真时拒绝 H_0；第二类错误是当 H_0 实际为假时接受 H_0。第一类错误是“以真为假”的错误，犯第一类错误的概率由 α 给出，α 越大，H_0 越容易错误地被拒绝；第二类错误是“以假当真”的错误，犯第二类错误的概率通常用 β 表示。由于犯这两类错误的概率之间存在着反比关系，因而，当样本数目 n 一定时，α 减小将使 β 增大。若希望同时减小犯两类错误的可能性，必须增加样本数目 n。

实际应用时，人们通常只能控制犯第一类错误的概率，也就是错误地拒绝 H_0 的概率，这个概率就叫做显著性水平。它一般在进行统计检验时事先给定。在选定 α 的大小时，应根据实际情况考虑。若宁可“以真为假”，则应把 α 取得小些，如 0.01；否则，α 可取大些。一般检验时，取 $\alpha = 0.05$，$\alpha = 0.01$ 较多。为了保证 β 不致太大，样本数目不宜太少（如至少不小于 5）。

（三）否定域

拒绝零假设 H_0 的区域称为否定域或拒绝域。否定域的大小

与显著性水平 α 的选取有关。对于同一组样本数据，选取不同的 α 值，可能得到截然相反的结论，如取 $\alpha = 0.05$，可能拒绝 H_0，而取 $\alpha = 0.01$，则不能拒绝 H_0。选择 α 的原则，不愿"以假当真"，而宁可"以真为假"时，应取很小的 α 值，正是基于否定域与 α 的关系建立的。

否定域的位置(不是大小)与备择假设 H_1 的性质有关。若 H_1 是指出预定方向的，如 $H_1: \mu > \mu_0$，则假设检验为单侧检验(单尾检验)；若 H_1 未指出预定方向，如 $H_1: \mu \neq \mu_0$，则为双侧检验(双尾检验)。图 1.1 是 $\alpha = 0.05$ 的单侧检验否定域，图 1.2 是 $\alpha = 0.05$ 的双侧检验否定域。可以看出，对于同一显著性水平 α，两种否定域的位置不同，但总的大小并没有什么不同。

在进行统计检验时，若根据样本数据计算的统计量数值落入否定域，则认为零假设 H_0 不成立，称作在显著性水平 α 下拒绝 H_0；否则认为零假设 H_0 成立，称作在显著性水平 α 下不能拒绝 H_0。



图 1.1 $\alpha = 0.05$ 的单侧检验否定域



图 1.2 $\alpha = 0.05$ 的双侧检验否定域

三、非参数统计方法

(一) 参数统计和非参数统计

在数理统计学中，统计检验的种类很多，而每一种统计检验都

第二章　单样本非参数检验

单样本非参数统计方法是用来检验只需抽取一个样本的假设。通常能回答下面的问题：观察频数和某种原则下的期望频数是否有显著差异；观察的比例与所期望的比例是否有显著差异；样本取自某种类型的总体的假定是否合理等等。单样本非参数检验通常属于拟合优度检验。

§2.1　χ^2 检验

χ^2 检验(Chi-Square Goodness-of-Fit Test)属于拟合优度检验，它可以用来检验样本内每一类别的实际观察数目与某种条件下的理论期望数目是否有显著差异。

一、基本方法

若一个事件只有两个可能的结果，如产品或合格或不合格，对某房改方案或赞成或反对，那么通常可以用参数检验的方法判定其观察频数是否显著地背离期望频数。但当一个事件可能有两个以上，如 k 个结果出现时，采用 χ^2 检验是最适合的。若样本分为 k 类，每类实际观察频数为 $f_1, f_2, \cdots, f_k$，与其相对应的期望频数为 $e_1, e_2, \cdots, e_k$，则统计量 Q 可以测度观察频数与期望频数之间的差异。其计算公式为：

$$Q = \sum_{i=1}^{k} \frac{(f_i - e_i)^2}{e_i} \tag{2.1}$$

很显然，观察频数与期望频数越接近，Q 值就越小，若 $Q = 0$，则(2.1)式中分子的每一项都必须是 0，这意味着 k 类中每一类观察频数与期望频数完全一样，即完全拟合。Q 统计量可以用来测度实际观察频数与理论期望频数之间的紧密程度即拟合程度。

若零假设为观察频数充分地接近期望频数，即对于 $i = 1, 2,$

…，k，f_i 与 e_i 无显著差异，则由于样本容量 n 充分大时，Q 统计量近似地服从自由度 $df = k - 1$ 的 χ^2 分布，因而，可以根据给定的显著性水平 α，在附表 I 中查到相应的临界值 $\chi^2_\alpha(k-1)$。若 $Q \geqslant \chi^2_\alpha(k-1)$，则拒绝 H_0，否则不能拒绝 H_0。

二、应用

χ^2 检验运用的领域很多，在单样本问题中大致可以用来解决下面几类问题。

1. 检验某个已知比例的假设

当假设总体的比例为某数值时，需要利用从总体中抽取的样本来检验这个假设是否成立，χ^2 检验是一种适用的方法。

【例 2.1】 某企业大批量连续生产某产品，要求不合格品率不大于 5%。现从产品总体中，抽取 100 个进行检查，不合格品有 12 个，试以 5% 的显著性水平检验该批产品的不合格品率是否为 5%。

分析：因为检验的是产品不合格品率是否为 5% 这一已知比例，因此是双尾检验。建立假设组为

$$H_0:\quad P = 0.05$$

$$H_1:\quad P \neq 0.05$$

在这批产品中期望的不合格品数为 $100 \times 0.05 = 5$，合格品的期望数为 95，即 $f_1 = 12$，$f_2 = 88$，$e_1 = 5$，$e_2 = 95$。于是由(2.1)式计算的 Q 统计量为：

$$\begin{aligned} Q &= (12-5)^2/5 + (88-95)^2/95 \\ &= 10.316 \end{aligned}$$

根据显著性水平 $\alpha = 0.05$，自由度 $df = k - 1 = 1$，查 χ^2 分布表（附表 I），得到 $\chi^2_{0.05} = 3.841$。由于 $Q = 10.316 > \chi^2_{0.05} = 3.841$，则拒绝 H_0。这表明在 5% 的显著性水平上，不能认为该批产品合格率为 95%，即不合格品率为 5%。

如果样本一开始分类就为两类，而其中一类的期望频数小于

5，那么最好不用 χ^2 检验而采用二项检验。

2. 检验某种已知比例的假设

很多的现象之间往往会表现为某种比例关系，例如对同一种疾病，不同药物治愈的比率，不同类型贷款的偿还比率等等，为了检验某种预期的比例是否成立，可以采用 χ^2 检验。

【例 2.2】 某金融机构的贷款偿还类型有 A、B、C、D 四种，各种的预期偿还率为 80%、12%、7% 和 1%。在一段时间的观察记录中，A 型按时偿还的有 380 笔，B 型有 69 笔，C 型有 43 笔，D 型有 8 笔。问在 5% 显著性水平上，这些结果与预期的是否一致。

分析：这个问题属于要检验每一类型的出现概率与预期概率是否相等，即

H_0: $P_i = P_{i0}$ 对于一切 $i = 1,2,\cdots,k$

H_1: $P_i \neq P_{i0}$ 对于一些 $i = 1,2,\cdots,k$

其中，$P_1 + P_2,\cdots + P_k = 1$

它仍可采用 χ^2 检验，通过实际观察频数与理论期望频数是否有显著差异作出判断。

H_0: $A:B:C:D$ 类型偿还贷款的标准比率为 80:12:7:1

H_1: 偿还贷款是一些其它比率

在观察的已偿还的 500 笔贷款中，A 的预期偿还数为 $500 \times 0.8 = 400$，其它的以此类推。表 2—1 给出了计算 Q 统计量的过程及结果。

表 2—1 Q 统计量计算表

类型	f_i	e_i	$f_i - e_i$	$(f_i - e_i)^2$	$(f_i - e_i)^2/e_i$
A	380	400	−20	400	1.00
B	69	60	9	81	1.35
C	43	35	8	64	1.83
D	8	5	3	9	1.80
合计	500	500	—	—	5.98

根据给定的显著性水平 $\alpha = 0.05$，自由度 $df = k - 1 = 4 - 1 = 3$，查 χ^2 分布表，得到 $\chi^2_{0.05} = 7.82$，由于

$$Q = 5.98 < \chi^2_{0.05} = 7.82$$

表明在 5% 的显著性水平上不能拒绝 H_0，即观察比例与期望比例很为一致。

χ^2 检验也可以这样来判定：根据自由度 df 和计算的 Q 值，在 χ^2 分布表上找到 H_0 成立时的概率 P，若 P 等于或小于 α，则拒绝 H_0，否则不能拒绝 H_0。例 2.2 中，根据 $df = 3$，$Q = 5.98$，查 χ^2 分布表，Q 值落在 $P = 0.20$ 和 $P = 0.10$ 所对应的 4.64 与 6.25 二者之间，由于 $P = 0.10 > \alpha = 0.05$，不能拒绝 H_0。

公式(2.1)中，由 $\Sigma f_i = n$，$\Sigma e_i = n$，n 为样本容量即各类观察频数之和，因而可以变形为：

$$\begin{aligned} Q &= \sum_{i=1}^{k} \frac{(f_i - e_i)^2}{e_i} = \frac{(f_1 - e_1)^2}{e_1} + \frac{(f_2 - e_2)^2}{e_2} + \\ &\quad \cdots + \frac{(f_k - e_k)^2}{e_k} = \frac{f_1^2}{e_1} + \frac{f_2^2}{e_2} + \cdots + \frac{f_k^2}{e_k} \\ &\quad + (e_1 + e_2 + \cdots + e_k) - 2(f_1 + f_2 + \cdots + f_k) \\ &= \sum_{i=1}^{k} f_i^2 / e_i - n \end{aligned} \qquad (2.2)$$

其中，$e_i = nP_{i0}$，P_{i0} 为理论期望概率。

【例 2.3】 续例 2.2　由于在例 2.2 中 D 类的期望频数等于 5，较小，则应与邻近的类别合并（必须保证合并是合理的），从而减小 k 值而增加某些类别的 e_i 值。若将 C 与 D 类合并，则 A、B、C 和 D 预期的概率分别为 0.8，0.12，0.08。根据观察的结果，能否在 5% 的显著水平上，作出结论。

分析：　H_0：　$P_i = P_{i0}$　　对于一切 $i = 1, 2, 3$

　　　　H_1：　$P_i \neq P_{i0}$　　对于一些 $i = 1, 2, 3$

Q 统计量利用(2.2)式计算。过程如表 2—2。

根据给定的显著性水平 $\alpha = 0.05$，$df = 2$，查 χ^2 分布表得到 $\chi^2_{0.05} = 5.99$。由于 $Q = 5.375 < \chi^2_{0.05} = 5.99$，表明在 5% 的显著性

水平上不能拒绝 H_0。

表 2—2　　　　Q 统计量计算表

类　型	f_i	P_{i0}	$e_i = nP_{i0}$	f_i^2	f_i^2/e_i
A	380	0.80	400	144400	361
B	69	0.12	60	4761	79.35
C 和 D	51	0.08	40	2601	65.025
合　计	500	1.00	500	—	505.375
	$Q = 505.375 - 500 = 5.375$			$\mathrm{df} = 3 - 1 = 2$	

有时，预期的比率不一定表现为概率形式，如四种药物对同一种疾病的有效比率为 9∶3∶3∶1。这时，可以将其转换为预期概率，即四种药物的期望有效率分别为$\frac{9}{16}$,$\frac{3}{16}$,$\frac{3}{16}$,$\frac{1}{16}$。

3. 检验总体是否为某一分布

如果随机抽取的样本其数值为 $x_1, x_2, \cdots, x_n$，来自概率密度函数未知的某一连续分布，将其总体的理论分布记作 $F_0(x)$，实际观察数据的分布为 $F(x)$，检验总体是否为某一特定分布 $F_0(x)$ 的假设为：

$$H_0:\quad F(x) = F_0(x) \qquad \text{对所有的 } x$$
$$H_1:\quad F(x) \neq F_0(x) \qquad \text{对一些 } x$$

利用 χ^2 检验可以作出判定。

将随机抽取的样本数据分为 k 组（要保证每组的期望频数至少是 5，否则应将相邻组合并），以样本参数估计值作为总体特定分布的参数值（因为是大样本，故以样本值替代总体值），计算每一组的期望频数 e_i，再以（2.1）式计算 Q 统计量。由于某一特定分布可能有 w 个参数是固定的，因而 Q 统计量的自由度 $\mathrm{df} = k - w - 1$。根据 α、df，查 χ^2 分布表，可以得到拒绝或不拒绝 H_0 的结论。表 2—3 是拟合优度检验中几种分布的参数。

表 2-3　　拟合优度检验中几种分布的参数

分　布	参　数	估计值	w	df
二项分布	一个试验成功的概率			
(n 次试验的)	θ	$\Sigma xf/n\Sigma f$	1	$k-2$
泊松分布	λ	$\bar{x}$	1	$k-2$
正态分布	μ,σ^2	$\bar{x},s^2$	2	$k-3$
指数分布				
$F(x)=1-e^{-\lambda x}$	$1/\lambda$	$1/\bar{x}$	1	$k-2$

【例 2.4】　两种不同牌号的茶哪个更好。

今有 30 人组成的品茶专家组，对 A、B 两种不同牌号的茶进行 6 种不同味道的检验。凡专家认为优者被记录下来，如表 2—4。不同牌号的茶提供给专家品尝是随机的。

分析：两种不同牌号的茶中，A 被选择是优的概率，可视为二项分布中一个试验成功的概率为多少的问题。由表 2—3 提供的方法，利用表 2—4 的数据，可以估计二项分布中的参数 θ，即一个试验成功的概率，这里就是 A 被选择为优的概率。因此，判断 A、B 茶哪个更好，实际上就是判断表 2—4 的样本数据是否来自参数 $\theta=\Sigma xf/n\Sigma f$ 的二项分布总体。若是，则哪个牌号被选为优的概率越大，哪个牌号的茶更好。

H_0：　$F(x)$ 为二项分布

H_1：　$F(x)$ 不是二项分布

由于对任何一个品尝专家来说，A 牌号都能被选择 1 至 6 次，因此，按 A 被选择的次数整理数据，得到表 2—5。其中，x 表示 A 被一个专家选择的次数，f 表示相应的专家人数，在这个问题中，对茶叶进行 6 个味道品尝检验，故 $n=6$，参数 θ 的估计值就是 $\Sigma xf/n\Sigma f=117/180=0.65$。根据 $n=6$，$\theta=0.65$，查附表 Ⅱ 可得到 P_0，利用 $e=NP_0(N=\Sigma f=30)$ 计算期望频数，可以将实际

频数 f 与期望频数 e 比较，采用 χ^2 检验判断观察频数与期望频数是否有显著差异，从而决定理论分布与实际分布是否一致，即是否服从二项分布。Q 统计量计算过程如表 2—6。

表 2-4　　专家检验茶叶味道结果统计表

专家	味道检验结果						A 被选择的次数
	1	2	3	4	5	6	
1	B	B	A	B	A	A	3
2	B	B	B	B	A	A	2
3	B	A	B	A	A	B	3
4	A	A	A	A	A	A	6
5	B	A	A	A	A	B	4
6	A	B	A	A	A	A	5
7	B	A	A	A	A	B	4
8	B	B	A	B	A	A	3
9	A	B	B	A	A	A	4
10	A	B	B	A	A	A	4
11	A	A	B	A	A	A	5
12	A	B	B	A	A	A	4
13	B	A	B	A	A	A	4
14	B	A	B	B	A	A	3
15	A	B	A	B	A	A	4
16	B	A	A	A	A	A	5
17	B	B	A	A	A	A	4
18	A	B	A	A	A	B	4
19	B	A	B	A	A	A	4
20	B	A	B	B	A	A	3
21	A	A	A	A	A	A	6
22	A	A	A	B	A	A	5
23	B	B	A	B	B	A	2
24	A	A	B	A	A	A	5
25	A	B	B	B	A	A	3
26	B	A	A	A	A	A	5
27	B	B	A	B	A	A	3
28	B	A	B	B	A	B	2
29	A	B	A	A	A	A	5
30	A	B	B	B	A	A	3

表 2—5　数据整理

A 被选择的次数(x)	频数(f)
0	0
1	0
2	3
3	8
4	10
5	7
6	2
合计	30

表 2—6 中，由于第一、二，三组的期望频数均小于 5，因此，将它们合并，计算结果 $Q = 0.2472$。根据显著性水平 $\alpha = 0.05$，自由度 $\mathrm{df} = k - w - 1 = 3$，查 χ^2 分布表得到 $\chi^2_{0.05} = 7.82$，因为 $Q = 0.2472 < \chi^2_{0.05} = 7.82$，所以不能拒绝 H_0，换句话说，从附表 I 可知，有 95% 以上的把握支持这批数据来自一个二项分布的假设，且参数 θ 估计值为 0.65。这一结果表明，A、B 两种牌号的茶叶中，A 被选择是优的概率较大，因此，专家评价结果 A 茶更好。

表 2—6　Q 统计量计算表

x	f	xf	P_0	$e = NP_0$	$f - e$	$(f-e)^2$	$(f-e)^2/e$
0	0	0 }	0.0018	0.054 }	−0.522		
1	0	0 } 6	0.0205	0.615 } 3.522		0.2725	0.0774
2	3	6 }	0.0951	2.853 }			
3	8	24	0.2355	7.065	0.935	0.8742	0.1237
4	10	40	0.3280	9.84	0.16	0.0256	0.0026
5	7	35	0.2437	7.311	−0.311	0.0967	0.0132
6	2	12	0.0754	2.262	−0.262	0.0686	0.0303
	30	117	1.0000	30			0.2472

§2.2　Kolmogorov-Smirnov 检验

Kolmogorov-Smirnov 检验简写为 K-S 检验，常译成柯尔莫哥洛夫—斯米尔诺夫检验。它也是一种拟合优度检验。它涉及一组样本数据的实际分布与某一指定的理论分布间相符合程度的问题，用来检验所获取的样本数据是否来自具有某一理论分布的总体。

一、基本方法

K-S 检验是用两个俄罗斯数学家的名字命名的，他们对这种非参数统计技术的发展作出了贡献。若 $S_n(x)$ 表示一个 n 次观察的随机样本观察值的累积概率分布函数，$S_n(x) = i/n$，i 是等于或小于 x 的所有观察结果的数目，$i = 1,2\cdots,n$。$F_0(x)$ 表示一个特定的累积概率分布函数，也就是说，对于任一 x 值，$F_0(x)$ 值代表小于或等于 x 值的那些预期结果所占的比例。于是，可以定义 $S_n(x)$ 与 $F_0(x)$ 之间差值，即

$$D = |S_n(x) - F_0(x)|$$

其中，$S_n(x)$ 为经验分布函数，$F_0(x)$ 为理论分布函数，若对每一个 x 值来说，$S_n(x)$ 与 $F_0(x)$ 十分接近，也就是差异很小，则表明经验分布函数与特定分布函数的拟合程度很高，有理由认为样本数据来自具有该理论分布的总体。K-S 检验集中考察的是 $|S_n(x) - F_0(x)|$ 中那个最大的偏差，即利用统计量

$$D = \max|S_n(x) - F_0(x)| \tag{2.3}$$

作出判定。

K-S 检验的步骤为：

建立假设：

H_0: $S_n(x) = F_0(x)$ 对所有 x

H_1: $S_n(x) \neq F_0(x)$ 对一些 x

计算 D 统计量：

$$D = \max|S_n(x) - F_0(x)|$$

查找临界值：根据给定的显著性水平 α，样本数据个数 n，查附表 Ⅲ 可以得到临界值 d_α（双尾检验）。

作出判定：若 $D < d_\alpha$，则在 α 的水平上，不能拒绝 H_0；若 $D \geq d_\alpha$，则在 α 水平上，拒绝 H_0。

二、应用

在许多实际问题中，检验确定某一组数据是否来自某一特定

$$C = \frac{1}{\sum_{x=1}^{19} 1/x^m + 1/(2 \cdot 19^m) + 1/[(m-1) \cdot 19^m + m]/(24 \cdot 19^{m+1})}$$

经过计算，$C = 1/(1.1908 + 0.000062 + 0.001146 + 0.0000008) \doteq 0.8389$。因此，46 期《数理统计与管理》的论文与作者数的理论洛特卡分布为

$$f(y) = 0.8389/x^{3.055} \tag{2.6}$$

为了判定《数理统计与管理》论文作者的实际分布是否与理论分布一致，可以采用 $K\text{-}S$ 检验。建立的假设组为

$$H_0:\quad S_n(x) = F_0(x) \qquad \text{对所有 } x$$

$$H_1:\quad S_n(x) \neq F_0(x) \qquad \text{对一些 } x$$

理论累积频率 $F_0(x)$ 的各个值，可以将 x 分别代入(2.6)式计算得到，实际累积频率是将累计的作者数 Σy 分别除以作者总人数得到。计算结果，作者实际累积频率及理论累积频率及各个差值如表 2—8。

表 2—8　　**作者实际累积频率与理论累积频率表**

x	1	2	3	4	5	6	7
$F_0(x)$	0.8389	0.9398	0.9690	0.9811	0.9872	0.9907	0.9929
$S_n(x)$	0.8932	0.9635	0.9870	0.9896	0.9922	0.9948	1.0000
$\|S_n(x) - F_0(x)\|$	0.0543	0.0237	0.0180	0.0085	0.0050	0.0041	0.0071

$$D = \max|S_n(x) - F_0(x)|$$
$$= 0.0543$$

根据显著性水平 $\alpha = 0.01$，作者人数 $n = \Sigma y = 384$，查附表 Ⅲ，得到临界值 $d_\alpha = 1.63/\sqrt{n} = 1.63/\sqrt{384} \doteq 0.0832$。由于 $n > 45$，为大样本，故应选取附表 Ⅲ 中最后一行的计算公式得到临界值。显然，

$$D = 0.0543 < d_\alpha = 0.0832$$

因此，数据在 1% 的显著性水平上不能拒绝 H_0，若显著性水平 α 取 0.05，查附表 Ⅲ，得到临界值 $d_\alpha = 1.36/\sqrt{n} = 1.36/\sqrt{384} =$

0.0694。显然

$$D = 0.0543 < d_\alpha = 0.0694$$

因此，数据在5%的显著性水平上也不能拒绝 H_0。可以认为，《数理统计与管理》作者的分布服从洛特卡分布。

【例2.6】 公共交通设施适合性的研究——公共汽车到达时间是否服从正态分布

公共汽车按计划每15分钟通过一个商店旁。然而，由于交通条件，乘客数目等的影响，汽车实际到达的时间有很大不同。通过一天随机的观察，获得的数据如表2—9。比计划提前到达的为负值，取大的整数，如提前1分10秒到达，记作 -1；比计划晚到的为正值，也取大的整数，如迟到1分10秒，记作 $+2$。公共汽车到达时间是否服从 $\sigma = 3$ 的正态分布。

表2—9　　汽车到达时间统计表

到达时间(x)	-5	-3	-1	0	1	2	4	7	8
观测频率(f)	1	1	2	1	5	5	3	1	1

分析：正态分布是一个常用的概率模型，如果公共汽车到达时间被证明是服从正态分布，就为进一步的研究提供了一个方便使用的模型。这里 $F_0(x)$ 是累积的正态分布函数，因为其是连续的，因此使用 K-S 检验是合适的。

$$H_0:\quad S_n(x) = F_0(x)\qquad 对所有\ x$$

$$H_1:\quad S_n(x) \neq F_0(x)\qquad 对一些\ x$$

这里 $F_0(x)$ 是标准正态分布 $\Phi(Z)$。

为了得到 $F_0(x) = \Phi(Z)$，要计算 μ，它的最好估计值是样本平均数 $\bar{x}$。$\bar{x} = \Sigma xf/\Sigma f = 1.6$。借助于 $Z = (x-\mu)/\sigma$ 将数据标准化，计算过程如表2—10。表中 Z 的概率一列，是根据 Z 的绝对值查找附表Ⅳ得到的。

$F_0(x) = \Phi(Z)$ 一列的数值是依据 Z 和 Z 的概率得到。若 Z 为负值，$\Phi(Z)$ 就是 Z 的概率，若 Z 为正值，则 $\Phi(Z)$ 是1减去 Z 的概

的值。正态概率模型对这一问题是适用的，但也许会有更好的其它概率模型。

表 2-11　　　　D 的计算表

x	$S_n(x)$	$F_0(x)$	$\|S_n(x)-F_0(x)\|$
-5	0.0500	0.0139	0.0361
-3	0.1000	0.0630	0.0370
-1	0.2000	0.1922	0.0078
0	0.2500	0.2981	0.0481
1	0.5000	0.4207	0.0793
2	0.7500	0.5517	0.1983
4	0.9000	0.7881	0.1119
7	0.9500	0.9641	0.0141
8	1.0000	0.9843	0.0166

三、χ^2 检验与 K-S 检验

χ^2 检验与 K-S 检验均属拟合优度检验，但 χ^2 检验常用于定类尺度测量数据，K-S 检验还用于定序尺度测量数据。当预期频数较小时，χ^2 检验常需合并邻近的类别才能计算，K-S 检验则不需要，因此它能比 χ^2 检验保留更多的信息。对于特别小的样本数目，χ^2 检验不能应用，而 K-S 检验则不受限制。因此，K-S 检验的功效比 χ^2 检验要更强。

§ 2.3　符号检验

符号检验(Sign Test)是利用正，负号的数目对某种假设作出判定的非参数统计方法。

一、普通的符号检验

1. 基本方法

如果所研究的问题，可以看作是只有两种可能："成功"或"失

败”，并且成功或失败的出现被假定遵从二项式分布，以 $+$ 表示成功，$-$ 表示失败，那么随机抽取的样本就有两个参数：成功的概率 P_+，失败的概率 P_-。这样，就可以构造一个假设：

$$H_0:\quad P_+ = P$$

$$H_1:\quad P_+ \neq P_-$$

这是双侧检验，对备择假设 H_1 来说，不要求 P_+ 是否大于 P_-。如果所研究的问题，要求考虑是 P_+ 比较大还是 P_- 比较大，则需用单侧备择假设，即

$$H_0:\quad P_+ = P_- \qquad\qquad H_0:\quad P_+ = P$$

$$H_+:\quad P_+ > P_- \qquad\qquad H_-:\quad P_+ < P_-$$

这里 H_+ 表示 P_+ 是比较大的，H_- 被用来说明 P_- 是比较大的。

为了检验上面的假设，普通的符号检验所定义的检验统计量为 S_+ 和 S_-。S_+ 表示为正符号的数目，S_- 表示为负符号的数目，$S_+ + S_- = n$，n 是符号的总数目。

要对假设作出判定，需要找到一个 P 值。因为对于 S_+ 和 S_- 来说，抽样分布是一个带有 $\theta = 0.5$（θ 表示成功的概率）的二项式分布，所以如果 H_0 为真，从附表 Ⅵ 中能够根据 n、S_+ 或 S_- 查到 P 值。若 P 值很小，表明 H_0 为真的可能性很小，数据不支持 H_0，而支持 H_1。

当样本的观察数据 $n \leqslant 20$ 时，可以利用上面方法找到 P 值作出判定。若样本的观察数据 $n > 20$，可以用正态近似办法，根据 (2.7) 式计算 Z 值，查找附表 Ⅳ 得到相应的 P 值。

$$Z_{+,R} = \frac{S_+ - 0.5 - 0.5n}{0.5\sqrt{n}}$$

$$Z_{-,R} = \frac{S_- - 0.5 - 0.5n}{0.5\sqrt{n}} \tag{2.7}$$

普通的符号检验其判定可以归纳如表 2—12。

2. 应用

在实际问题的研究中，常常会遇到难以用数值确切表达的问

啡、茶中更喜欢茶为“成功”，反之为“失败”故可建立假设

H_0:　$P_+ = P_-$

H_1:　$P_+ \neq P_-$

这一假设仅判定对二者喜欢程度有无差异。由调查结果知：$S_+ = 12$，$S_- = 2$，$n = 14$。查附表 Ⅵ，$n = 14$ 时，S_+ 与 S_- 中大者 $S_+ = 12$ 右尾概率的 2 倍是 $2 \times 0.0065 = 0.0130$。显然 $P = 0.0130$ 很小，以显著性水平 $\alpha = 0.05$，P 足够小，故这批数据不支持 H_0，即顾客对咖啡和茶的喜爱有显著差异。判断是否更喜欢茶，建立单侧备择假设

H_0:　$P_+ = P_-$

H_+:　$P_+ > P_-$

根据上面的分析，$n = 14$，$S_+ = 12$ 时，附表 Ⅵ 显示的 P 值为 0.0065。对于 $\alpha = 0.01$，P 值也足够小，故数据不支持 H_0，而支持 H_+。顾客在咖啡和茶中更喜欢茶。

二、位置的符号检验

1. 基本方法

一个随机抽取的样本，有 n 个数据 $x_1, x_2, \cdots, x_n$，其实际的总体中位数记作 M，假定的中位数是某个特定值，记作 M_0。若研究问题时关心的是：真实的中位数 M 是否不同于 M_0，可以建立假设

H_0:　$M = M_0$

H_1:　$M \neq M_0$

如果关心的是：真实的中位数 M 是否大于或小于特定的数 M_0，则应建立单侧备择假设

H_0:　$M = M_0$　　　H_0:　$M = M_0$

H_+:　$M > M_0$　　　H_-:　$M < M_0$

只要样本数据 $x_1, x_2, \cdots, x_n$ 能够被测量，至少是定距尺度测量，若定序尺度测量的，应能与 M_0 相比并决定大小次序，也就是说，数据与 M_0 比较的结果能用“+”或“-”表示，设定 $x_1, x_2, \cdots, x_n$ 在 M_0

附近是连续的，则普通的符号检验方法可以应用。这就是符号检验法用于单样本位置的推断。

检验统计量。位置的符号检验所定义的检验统计量也是S_+和S_-，S_+表示每一个观察数据$x_i(i=1,2\cdots,n)$与特定的数M_0的差值$D_i=x_i-M_0$的符号为正的数目，S_-表示D_i符号为负的数目。$S_++S_-=n$，n表示符号的总数目。

P值的确定。当$n\leqslant 20$时，与普通的符号检验一样，从附表Ⅵ中查找；若$n>20$，由正态近似得到P值，在附表Ⅳ中查找。

2. 应用

【例2.10】 生产过程是否需要调整

某企业生产一种钢管，规定长度的中位数是10米。现随机地从正在生产的生产线上选取10根进行测量，结果为：

9.8，10.1，9.7，9.9，9.8，10.0，9.7，10.0，9.9，9.8。

分析：中位数是这个问题中所关心的一个位置参数。若产品长度真正的中位数大于或小于10米，则生产过程需要调整。这是一个双侧检验，应建立假设

$$H_0:\quad M=M_0=10$$

$$H_1:\quad M\neq M_0\neq 10$$

为了对假设作出判定，先要得到检验统计量S_+，S_-。将调查得到数据分别与10比较，算出各个符号的数目：$S_+=1$，$S_-=7$，$n=8$（差值$D_i=0$的不计）。在附表Ⅵ中，$n=8$，S_+与S_-中较大者7的右尾概率的2倍是$2\times 0.0352=0.0704$。若显著性水平$\alpha=0.05$，P已足够大，表明调查数据支持H_0。根据这批数据，生产过程暂不需做调整。

3. 配对样本位置的符号检验

作为单样本位置的符号检验，也能够应用于配对样本。随机选取的n个数据对(x_1,y_1)，(x_2,y_2)，$\cdots$，(x_n,y_n)至少是定距尺度测量，若是定序尺度测定，那么就要既对于样本的数据，又要对于M_0都是定序尺度。因为，为利用符号检验，至少要求每对数据的差值

$D_i = x_i - y_i$ 相对于 M_0 来说都是可以测定的。假定差值总体在 M_0 附近是连续的，M_D 表示随机变量的数对之间差值 $D_1, D_2, \cdots, D_n$ 的总体中位数，M_0 仍表示特定的数，则将位置的符号检验假设中的 M 以 M_D 替代就可以得到配对样本的假设组

$$H_0: \quad M_D = M_0$$

$$H_1: \quad M_D \neq M_0$$

$$H_0: \quad M_D = M_0 \qquad H_0: \quad M_D = M_0$$

$$H_+: \quad M_D > M_0 \qquad H_-: \quad M_D < M_0$$

配对样本的方法同单样本位置的符号检验方法一样，检验统计量也是 S_+，S_-。但是，S_+ 是差值 D_i 与 M_0 的差值的符号为正的数目，S_- 是 $D_i - M_0$ 为负的符号数目。$D_i - M_0 = x_i - y_i - M_0$。和以前一样，若 $D_i - M_0 = 0$，则 n 中不计。判定假设的 P 值仍在附表 Ⅵ 中，按以前那样查找。若 $n > 20$，也是利用 S_+、S_- 按(2.5)式计算 Z，在附表 Ⅳ 中查找相应的 P 值。

【例 2.11】 领导者的领导水平是可以训练的

为验证领导水平是可以训练的，根据人的聪明程度、人品、受教育状况等，随机抽选出 12 个人配成 6 对，每对中有一人随机选择受训，记作 T，另一人则不受训记作 C。经过一段时间后，按被设计好的问题评价他们的领导水平，结果如表 2—14。

表 2—14　　配对样本评价表

配对样本编号	T 的评价(x)	C 的评价(y)	$x - y$
1	13	10	+
2	19	7	+
3	34	20	+
4	24	38	—
5	40	22	+
6	39	15	+

分析：表中的分数并不是精确测量的结果，它只是相对的领

导能力比较。它没有一个精确的量的意义，而是一对中两人比较，反映领导水平差异性的符号。如果领导水平是可以训练的，那么接受特定训练的成员应该比那些配对的非受训人员有较高的分数，因此备择是单侧的。合适的假设为

$$H_0: \quad M_D = 0$$

$$H_+: \quad M_D > 0$$

由表 2—14 给出的数据可知，$S_+ = 5$，$S_- = 1$，$n = 6$，在附表 Ⅵ 中查找的准确 P 值为 $n = 6$，$S_+ = 5$ 的右尾概率，即 $P = 0.1094$。对于显著性水平 $\alpha = 0.05$，显然 P 是够大的。因此这批数据支持 H_0，不支持 H_+，即这批数据不足以说明领导水平是可以训练的。

【例 2.12】 人的智商是否随出生顺序而升高

通过对以往在校学生的调查发现，智商往往随着出生顺序而升高，这种关系在成年人中是否也存在?通过随机选取 13 对成人的同胞兄弟姐妹进行调查，得到智商的数值如表 2—15。

表 2—15　　配对样本智商值统计表

第一胎出生的智商值(x)	以后出生的智商值(y)	($x-y$) 的符号
86	82	+
90	94	—
91	96	—
101	106	—
93	92	+
85	90	—
92	98	—
115	120	—
72	74	—
75	80	—
120	130	—
106	110	—
104	109	—

分析：如果成年人的智商与出生顺序有关，并与孩子的规律相同，那么头一胎出生的人的智商值应该比自己的其他兄弟姐妹低。因此应建立单侧备择假设。差值 $D_i = x_i - y_i$ 是第一胎的智商

值减去后面出生的弟或妹的智商值。对于零假设，应该是他们的智商值无差异，也就是智商值差值的中位数 $M_0 = 0$。这是一个配对样本位置的符号检验。建立在差值中位数基础上的假设为

$$H_0:\quad M_D = 0$$

$$H_-:\quad M_D < 0$$

由表 2—15 可知，$S_+ = 2$，$S_- = 11$，$n = 13$。在附表 Ⅵ 中相对于 $n = 13$，$S_- = 11$ 的右尾概率是 0.0112，这是准确的 P 值。对于显著性水平 $\alpha = 0.05$，P 值要小得多。这意味着，对于那些差值的中位数为 0 的随机变量，在 13 个符号中，有 11 个是负号的情况极少出现。因而数据不支持 H_0，而支持 H_-。这表明对于成年人来说，出生顺序仍然对智商有影响，智商随着出生顺序而增高。

三、中位数的置信区间

符号检验的统计方法可以用来估计单样本总体中位数的置信区间、配对样本差值总体中位数的置信区间。

1. 顺序统计量

顺序统计量(Order Statistic)亦称次序统计量。它在非参数统计中有很多的应用。假设 $X_1, X_2, \cdots, X_n$ 是抽自连续总体的一组随机变量样本数据，按其大小顺序由小到大依次排列，并引入一个带括号的下标表示其排列的位置，如 $X_{(r)}$ 表示在 n 个数据中的第 r 个，则它是来自于从最小开始排列的第 r 个数。类似的，可以将样本数据 $X_1, X_2, \cdots, X_n$ 依次排列为

$$X_{(1)} < X_{(2)} < \cdots < X_{(n)}$$

这就是顺序统计量，是原有样本数据按逐渐升高的顺序排列的。这里，若 $r < s$，那么 $X_{(r)} < X_{(s)}$ 只是表明它们以什么样的顺序被观察，按从小到大的顺序 $X_{(r)}$ 在前，$X_{(s)}$ 在后，至于它们的相对数值大小是多少并没有意义。顺序统计量没有与原有变量相同的概率函数，即使原变量是独立的，顺序统计量事实上可能不独立。

2. 置信区间

即

$$X_{(3)} \leqslant M \leqslant X_{(8)}$$

在顺序统计量中，处于第 3 个位置的数值是 9.8，处于第 8 个位置的是 10.0，因此，中位数的 89.06% 置信区间是

$$9.8 \leqslant M \leqslant 10.0$$

也就是说，可以有近 90% 的把握，钢管的中位数处在 9.8 米至 10 米之间。

对于单侧检验来说，置信区间估计的方法基本同上，只是对于置信系数 γ 来说，在附表 Ⅵ 中查找 k 时，P 不是 $(1-\gamma)/2$，而是 $1-\gamma$。例如对应于单侧检验 $H_-: M < M_0$ 估计 $M \geqslant X_{(k+1)}$ 的 95% 置信区间，根据 $r = 0.95, 1-r = 0.05, n = 10$，查附表 Ⅵ，得到与 $P = 0.0547$ 相应的，S_+ 的左尾数目为 2，即与 $\gamma = 0.9453$ 对应的 $k = 2$。因此，钢管中位数 $M \geqslant 9.8$ 的置信系数为 94.53%。对于单侧检验 $H_+: M > M_0$ 来说，估计 $M \leqslant X_{(n-k)}$ 的 95% 置信区间，在附表 Ⅵ 中，根据 $P = 1-\gamma = 0.05, n = 10$。查找，由于在表 2—12 中与单侧检验 H_+ 相对应的是 S_- 的左尾概率，所以 $k = 2$，即 M 的 94.53% 的置信区间为 $M \leqslant X_{(10-2)}$，也就是 $M \leqslant 10.0$。

四、二项式检验和分位数检验

1. 二项式检验

在实际问题中，有许多总体可以划分为两类，如将人分为男性和女性，学者和非学者；对患同一种病的人某种药物起作用还是不起作用等等，这是可以分作两类的总体，也称为二项总体。普通的符号检验可以用于来自任何二项总体的样本数据。若 θ_0 是一个指定的数值，$0 < \theta_0 < 1$，那么可以采用符号检验，判定某一些数据是否来自"成功"概率 $\theta = \theta_0$ 的二项总体。建立的双侧假设检验为

$$H_0: \quad \theta = \theta_0$$

$$H_1: \quad \theta \neq \theta_0$$

单侧检验为

$$H_0: \theta = \theta_0 \qquad H_0: \theta = \theta_0$$
$$H_+: \theta > \theta_0 \qquad H_-: \theta < \theta_0$$

随机抽取的样本数据个数为 n，或 n 次独立试验，或是 n 对相互比较的数组，都可以考虑应用符号检验判定是否来自带有参数 θ_0 的二项总体。在这 n 个数据中，每次观察都被分为成功或失败，作为成功的概率是 θ_0。S_+ 表示成功的数目，S_- 表示失败的数目。在 H_0 为真时，成功的期望数目是 $n\theta_0$，失败的数目是 $n(1-\theta_0)$。S_+ 是遵从带有参数 θ_0 的二项分布，S_- 是带有参数 $1-\theta_0$ 的二项分布（注意：$\theta_0 = 1-\theta_0$ 只有且仅当 $\theta_0 = 0.5$ 时成立，通常情况下，S_+ 和 S_- 不是同样分布）。

S_+ 和 S_- 被作为检验统计量。对于任何的 θ_0，当 S_+ 比它期望的值 $n\theta_0$ 大得多时，H_+ 被支持；若 S_+ 远远地小于 $n\theta_0$ 时，则 H_- 被支持。对于不同的备择假设，可以选择不同的检验统计量。将其总结如表 2—16。本书中附表 Ⅱ 是左尾概率，因而常用表 2—17 提供的检验统计量查找。

表 2—16　　二项式检验判定指导表

备择假设	P 值 （附表 Ⅱ）
$H_+: \theta > \theta_0$	带有 θ_0 的 S_+ 右尾概率
$H_-: \theta < \theta_0$	带有 θ_0 的 S_+ 的左尾概率
$H_1: \theta \neq \theta_0$	带有 θ_0 的 S_+ 小尾概率的 2 倍
$H_+: \theta > \theta_0$	带有 $1-\theta_0$ 的 S_- 的左尾概率
$H_-: \theta < \theta_0$	带有 $1-\theta_0$ 的 S_- 的右尾概率
$H_1: \theta \neq \theta_0$	带有 $1-\theta_0$ 的 S_- 小尾概率的 2 倍

以上是当 $n \leqslant 20$ 时的应用。$n > 20$ 时，可以采用正态近似解决，用(2.9)式先计算 $Z_{+,R}$，$Z_{-,R}$ 统计量，再查附表 Ⅳ，作出判定。表 2—18 是当 $n > 20$ 时的判定指导表。

表 2-17　　二项式检验判定指导表

备择假设	P 值　（附表 Ⅱ）
H_+：$\theta > \theta_0$	带有 $1-\theta_0$ 的 S_- 左尾概率
H_-：$\theta < \theta_0$	带有 θ_0 的 S_+ 的左尾概率
H_1：$\theta \neq \theta_0$	带有 θ_0 的 S_+，带有 $1-\theta_0$ 的 S_- 小尾概率的 2 倍

$$Z_{+,R} = \frac{S_+ - 0.5 - n\theta_0}{\sqrt{n\theta_0(1-\theta_0)}} \quad Z_{-,R} = \frac{S_- - 0.5 - n(1-\theta_0)}{\sqrt{n\theta_0(1-\theta_0)}} \tag{2.9}$$

表 2-18　　二项式检验判定指导表

备择假设	P 值　（附表 Ⅳ）
H_+：$\theta > \theta_0$	$Z_{+,R}$ 右尾概率
H_-：$\theta < \theta_0$	$Z_{-,R}$ 右尾概率
H_1：$\theta \neq \theta_0$	$Z_{+,R}$ 和 $Z_{-,R}$ 大者右尾概率的 2 倍

【例 2.14】　商场晚上是否应延长营业

某商场每晚 6：30 关门。有人建议应延长营业时间至 10：00。为作出决定，现欲对商场周围顾客情况作一调查，若商场的经常性顾客有 25% 以上说延长营业将去购买商品，则延长营业时间值得。随机选取 50 个家庭，发现只有 18 个能够被认为是商场的经常性顾客。调查结果，有 7 个家庭表示，延长营业时间将可能去购买。

分析：这个问题可以看作一个两分类总体，延长营业顾客去商场购买定义为成功。需要调查数据支持这样一个结论：成功的概率超过 25%。因此，合适的假设组为

$$H_0: \quad \theta = 0.25$$
$$H_+: \quad \theta > 0.25$$

由调查的结果可知，$S_+ = 7$，$S_- = 11$，$n = 18$。查附表 Ⅱ，$\theta = 1 - \theta_0 = 1 - 0.25 = 0.75$，$S_- = 11$，$P$ 值为 0.1390。这一结果表明，对显著水平 $\alpha = 0.05$，P 值是足够大的，即 $\theta = 0.25$ 不是小概率事件，

而是有足够大的概率发生。因此，数据支持 H_0，也就是这些调查数据不足以支持延长营业时间的假设。

2. 分位数检验

对于定序尺度测量的数据，描述其分散程度的指标是分位数。常用的有百分位数，十分位数和四分位数。将顺序统计量分为四等分，每一等分内含有样本总数的 25% 个数据，这些分割点称为第一个，第二个、第三个四分位点，与之相应的数值分别是第一个、第二个、第三个四分位数，以 Q 表示。将顺序统计量分为十等分，每等分内包含样本总数的 10% 个样本数据，分割点称为第一个、第二个、…、第九个十分位点，与之相对应的数值就是第一个、第二个、…、第九个十分位数，也可以 Q 表示。将顺序统计量分为一百等分，每等分内包含样本数据总数的 1% 个数据，这些分割点称为第一个、第二个、…，第九十九个百分位点，与之相对应的数值分别是第一个、第二个、…，第九十九个百分位数，也常以 Q 表示。实际应用时，为表示是第几个分位数，常给 Q 以右下标，如 Q_1 表示第一个分位数。符号检验的方法可以用于各种分位数的检验及置信区间估计。

随机抽取的样本，其 n 个样本数据，或 n 对数据的差值，按从小到大的顺序排列，得到一个顺序统计量。其某一分位数记作 Q，Q_0 是与之相应的一个特定的值，那么 Q 是否等于 Q_0？二者关系如何？为此，可以进行双侧检验或单侧检验，所建立的假设组为

$$H_0:\quad Q = Q_0$$
$$H_1:\quad Q \neq Q_0$$

$$H_0:\quad Q = Q_0 \qquad H_0:\quad Q = Q_0$$
$$H_+:\quad Q > Q_0 \qquad H_-:\quad Q < Q_0$$

要对假设作出判定，采用的检验统计量仍旧是 S_+、S_-。S_+ 是顺序统计量中数值大于 Q_0 的个数，S_- 是小于 Q_0 的数目（等于 Q_0 被忽略不计，这时 n 将减小）。对于一个分位数 Q，若在 Q 之下的概率是 P_-，在 H_0 为真时，S_- 是遵从带有参数 P_- 的二项分布，S_+ 则是遵

从带有参数 $P_+ = 1 - P_-$ 的二项分布。若以 θ_0 表示 P_+，则 P_- 就是 $1 - \theta_0$。与二项式检验类似，S_+ 遵从带有参数 θ_0 的二项分布，S_- 遵从带有参数 $1 - \theta_0$ 的二项分布。因而，当 $n \leqslant 20$ 时，查找附表 Ⅱ，可以得到相应的 P 值，从而作出判定。表 2—19 是判定指导表。当 $n > 20$ 时，可以采用正态近似，按(2.10)式计算 $Z_{+,R}$、$Z_{-,R}$ 查附表 Ⅳ，得到 P 值作出判定。表 2—20 是它的判定指导表。

表 2—19　　**分位数检验判定指导表**

备择假设	P 值　（附表 Ⅱ）
H_+：　$Q > Q_0$	带有 $1 - \theta_0$ 的 S_- 左尾概率
H_-：　$Q < Q_0$	带有 θ_0 的 S_+ 的左尾概率
H_1：　$Q \neq Q_0$	带有 θ_0 的 S_+，带有 $1 - \theta_0$ 的 S_- 的较小的左尾概率的 2 倍

$$Z_{+,R} = \frac{S_+ - 0.5 - n\theta_0}{\sqrt{n\theta_0(1 - \theta_0)}} \quad Z_{-,R} = \frac{S_- - 0.5 - n(1 - \theta_0)}{\sqrt{n\theta_0(1 - \theta_0)}} \tag{2.10}$$

表 2—20　　**分位数检验($n > 20$)判定指导表**

备择假设	P 值　（附表 Ⅳ）
H_+：　$Q > Q_0$	$Z_{+,R}$ 右尾概率
H_-：　$Q < Q_0$	$Z_{-,R}$ 右尾概率
H_1：　$Q \neq Q_0$	$Z_{+,R}$ 和 $Z_{-,R}$ 较大者右尾概率的 2 倍

【例 2.15】　例 2.12 中智商差值的第一个四分位数小于 -3。

分析：为判定第一胎出生和以后出生的成年人智商差值的第一个四分位数是否小于 -3，实际上就是要建立关于四分位数的单侧假设，其假设组为

$$H_0: \quad Q = -3$$
$$H_-: \quad Q < -3$$

这是一个配对样本差值的分位数检验。表 2—21 是根据表 2—15 数据计算的各对数据的差值，以及差值与 $Q_0 = -3$ 之差的符号。由表 2—21 可知，$S_+ = 3$，$S_- = 10$，$n = 13$。第一个四分位数 Q 表示比其小的数据个数是顺序统计量中样本数据总数的 25%，即在其之下的概率 $P = 0.25$。因此，$\theta_0 = P_+ = 1 - P_- = 0.75$。在附表 Ⅰ 中，查 $n = 13$，$\theta = 0.75$，$S_+ = 3$ 的左尾概率 $P = 0.0001$。这个 P 值对显著性水平 $\alpha = 0.05$ 来说足够小，，因此，调查数据不支持 H_0，表明第一胎出生的智商与以后出生的智商之差小于 -3 的数目，超过调查对数的 25%。

表 2—21

x	y	$x-y$	$x-y-(-3)$ 的符号
86	82	4	+
90	94	−4	−
91	96	−5	−
101	106	−5	−
93	92	1	+
85	90	−5	−
92	98	−6	−
115	120	−5	−
72	74	−2	+
75	80	−5	−
120	130	−10	−
106	110	−4	−
104	109	−5	−

【例 2.16】 今天成年人的睡眠量是否少于 5 年前

5 年前某地区的调查表明，成年人在每日 24 小时中的睡眠量中位数是 7.5 小时，每日睡眠量为 6 小时或少于 6 小时的占调查总数的 5%，9 小时和 9 小时以上的也占 5%。现对 8 个普通成年人的抽样调查结果为，每日 24 小时中睡眠量分为 7.2，8.3，5.6，7.4，7.8，5.2，9.1 和 5.8。

分析：为利用这批调查数据考察如今成年人睡眠量是否比 5

年前减少，可以采用分位数检验。根据5年前调查的数据，对于0.05，0.50，0.95的分位数，应该至少检验一个假定，即进行百分位数的检验。对中位数是否减少的检验即是对第0.50个分位数的检验。建立假设组

$$H_0: Q_{0.50} = 7.5$$

$$H_-: Q_{0.50} < 7.5$$

根据调查数据可知，$S_- = 3, S_+ = 5, n = 8$。由于$\theta_0 = 0.50$，因而可查附表Ⅵ($\theta = 0.50$)，也可查附表Ⅱ(在不同的θ值中，找到$\theta = 0.50$的那一列)。附表Ⅱ中，$n = 8, \theta_0 = 0.50, S_+ = 3$的左尾概率$P = 0.3633$。对于显著性水平$\alpha = 0.05$，$P$值已足够大，调查数据支持$H_0$，表明如今成年人睡眠量中位数与5年前无显著差异。

对第0.05个分位数即第5个百分位数进行检验，考察如今睡眠量是6小时和少于6小时的人数是否超过5%，即第0.05个分位数是否小于6，建立假设组

$$H_0: \quad Q_{0.05}{}^{①} = 6$$

$$H_-: \quad Q_{0.05} < 6$$

根据调查结果可知，$S_+ = 5, S_- = 3, n = 8$。在附表Ⅱ中，依据$\theta_0 = 1 - P_- = 1 - 0.05 = 0.95, n = 8, S_+ = 5$，查得$P = 0.0058$。对于显著性水平$\alpha = 0.01$，$P$已足够小，数据不支持$H_0$，表明如今睡眠量构成的顺序统计量中，第5个百分位数已小于6小时，也就是说，如今成年人睡眠量是6小时和少于6小时的人数已超过总人数的5%。

对第95个百分位数进行检验，建立假设组

$$H_0: \quad Q_{0.95} = 9$$

$$H_-: \quad Q_{0.95} < 9$$

根据调查数据可知，$S_+ = 1, S_- = 7, n = 8$。按表2—19提供的判定指导准则，在附表Ⅱ中，以$\theta_0 = 1 - P_- = 1 - 0.95 = 0.05$，

①$Q_{0.05}$表明这是第5个百分位数。

单侧备择假设

$$H_0:\quad M = M_0 \qquad H_0:\quad M = M_0$$

$$H_+:\quad M > M_0 \qquad H_-:\quad M < M_0$$

为了对假设作出判定，需要从总体中随机抽取一个样本得到 n 个观察值。这 n 个数据至少是定距尺度测量，若是定序尺度测量，则检验所需的等级、符号都应能被得到。n 个观察值记作 $x_1, x_2, \cdots, x_n$，它们分别与 M_0 的差值为 D_i，$D_i = x_i - M_0 (i = 1, 2, \cdots, n)$。如果 H_0 为真，那么观察值围绕 M_0 分布，即 D_i 关于 0 对称分布。这时，对于 D_i 来说，正的差值和负的差值应近似地相等。为了借助等级大小作判定，先忽略 D_i 的符号，而取绝对值 $|D_i|$。对 $|D_i|$ 按大小顺序分等级，等级 1 是最小的 $|D_i|$，等级 2 是第二小的 $|D_i|$，以此类推，等级 n 是最大的 $|D_i|$。按 D_i 本身符号的正、负分别加总它们的等级即秩次，得到正等级的总和与负等级的总和。虽然等级本身都是正的，但这里是按 D_i 符号计算的等级和。为了区别，将 D_i 符号为正的，$|D_i|$ 的等级和称作正等级的总和，反之为负等级的总和。H_0 为真时，正等级的总和与负等级的总和应该近似相等。如果正等级的总和远远大于负等级的总和，表明大部分大的等级是正的差值，即 D_i 为正的等级大。这时，数据支持备择假设 $H_+: M > M_0$。类似的，如果负等级的总和远远大于正等级的总和，表明大部分大的等级是负的差值，即 D_i 为负的等级大。这时，数据支持备择假设 $H_-: M < M_0$ 因为正等级和负等级的总和是个恒定的值，即

$$1 + 2 + \cdots + n = n(n+1)/2$$

因此对于双侧备择 H_1 来说，两个总和中无论哪一个太大，都可以被支持。

检验统计量。Wilcoxon 符号秩检验所定义的检验统计量为 T_+ 和 T_-。

T_-：　正等级的总和即正秩次总和

T ：　负等级的总和即负秩次总和

这里，T_+ 和 T_- 都是非负的整数，并且

$$T_+ + T_- = n(n+1)/2$$

它们的取值范围是从 0 到 $n(n+1)/2$。

P 值的确定。由于T_-和T_-的对称性，加上$T_+ + T_- = n(n+1)/2$，因而，T_+ 和 T_- 的抽样分布完全一样，且关于$n(n+1)/4$对称。附表 Ⅶ 给出了一个累积的概率，根据 n，查 T_+ 的右尾概率或 T_- 的右尾概率，得到 P 值。依据 P 与显著性水平 α 比较，可以对数据是否支持 H_0 作出判定。表 2—22 是 $n \leqslant 15$ 时判定指导表。

表 2—22　Wilcoxon 符号秩检验判定指导表

备择假设	P 值　（附表 Ⅶ）
H_+：　$M > M_0$	T_+ 的右尾概率
H_-：　$M < M_0$	T_- 的右尾概率
H_-：　$M \neq M_0$	T_+ 和 T_- 大者右尾概率的 2 倍

当 n 很大时，T_+、T_- 的标准化值近似于正态分布。T_+、T_- 的标准化可以借助于减去均值 $n(n+1)/4$，除以标准差 $\sqrt{n(n+1)(2n+1)/24}$ 做到。按(2.11)式计算 $Z_{+,R}$，$Z_{-,R}$，查附表 Ⅳ，可以得到相应的 P 值。表 2—23 是 $n > 15$ 时的判定指导表。

$$Z_{+,R} = \frac{T_+ - 0.5 - n(n+1)/4}{\sqrt{n(n+1)(2n+1)/24}}$$
$$Z_{-,R} = \frac{T_- - 0.5 - n(n+1)/4}{\sqrt{n(n+1)(2n+1)/24}} \tag{2.11}$$

表 2—23　Wilcoxon 符号秩检验判定指导表

备择假设	P 值　（附表 Ⅳ）
H_+：　$M > M_0$	$Z_{+,R}$ 的右尾概率
H_-：　$M < M_0$	$Z_{-,R}$ 的右尾概率
H_1：　$M \neq M_0$	$Z_{+,R}$ 和 $Z_{-,R}$ 大者右尾概率的 2 倍

Wilcoxon 符号秩检验也可以用来检验关于总体平均数的假

$-M_0$。

【例 2.18】 新配方是否有助于防晒黑

某防晒美容霜制造者，欲了解一种新配方是否有助于防晒黑，对 7 个志愿者进行了试验。在每人脊椎一侧涂原配方的美容霜，另一侧涂新配方的美容霜。背部在太阳下暴晒后，按预先给定的标准测定晒黑程度如表 2—25。

表 2—25　两种配方晒黑程度测量结果

编号	原配方(y)	新配方(x)	编号	原配方(y)	新配方(x)
1	42	38	5	44	33
2	51	53	6	55	49
3	31	36	7	48	36
4	61	52			

分析：这一问题是配对样本，同一个人使用两种不同配方的美容霜，以检验两种配方对防晒黑的作用是否有显著不同，可以应用 Wilcoxon 符号秩检验。新配方的晒黑程度记作 x，原配方的为 y，则差值 $d_i = x_i - y_i$。假定 d_i 关于中位数对称，那么两种配方的作用无显著差异时，d_i 的总体中位数应是0，即 $M_0 = 0$.为检验新配方是否优于原配方，则应建立的假设组为

$$H_0: \quad M_D = 0$$

$$H_-: \quad M_D < 0$$

为对假设作出判定，要计算 T_+、T_-，计算过程如表 2—26。由表可知

$$T_+ = 1 + 3 = 4$$

$$T_- = n(n+1)/2 - T_+ = 7(7+1)/2 - 4 = 24$$

根据 $n = 7, T_- = 24$ 查附表 Ⅶ，得到 T_- 右尾概率 $P = 0.055$。这个概率对于显著性水平 $\alpha = 0.05$ 来说已够太。一般在样本较小时，宜取稍大些的显著性水平，以避免犯第二类的错误，即“以假为真”

的错误。取显著性水平 $\alpha = 0.05$，则该调查数据支持 H_0，即两种配方对防晒黑的作用没有什么显著差异。若对于显著性水平 $\alpha = 0.10$，该次调查结果则不支持 H_0，即调查足以说明新配方对防晒黑的效果优于原配方。若希望同时减小犯两类错误的可能性，应增加样本数目。

表 2－26　　　　检验统计量计算表①

编　号	x	y	$D=x-y$	$\lvert D\rvert$	$\lvert D\rvert$ 的秩	D 的符号
1	38	42	−4	4	2	−
2	53	51	2	2	1	+
3	36	31	5	5	3	+
4	52	61	−9	9	5	−
5	33	44	−11	11	6	−
6	49	55	−6	6	4	−
7	36	48	−12	12	7	−

二、符号检验和 Wilcoxon 符号秩检验比较

1. 共同点

符号检验和符号秩检验都是非参数检验，都能运用于单一观察的数据或配对观察数据的差，都能用于总体中位数或差值总体中位数的推断。它们对总体所要求的假定都是极小的；对符号检验来说，是总体连续；对符号秩检验来说，再增加一个关于中位数对称。这两种检验数据测量层次的要求都不高。普通的符号检验被使用于两分类总体，类似于回答“是”或“不是”的问题，可用于定类、尺度测量，但要求差异的方向能够被表示出；符号秩检验至少要求定序尺度测量，仅当等级和符号能够被表示出时。由于两个检验都与符号有关，因而处理 0 差值的方法是共同的，均被忽略不计。

2. 区别

① 表中 $D=x\text{-}y$ 是因为 $M_0=0$。

符号检验仅使用各个观察值和中位数差值或配对样本差值与中位数差值方向上的信息，而没有考虑差值的大小，符号秩检验不仅利用差值方向上的信息，还利用了差值大小的信息，因此，它提供的信息量要多于符号检验。关于两种检验的功效有过不少的研究和报道，有兴趣的读者可以去阅读有关书籍。在大多数情况下Wilcoxon 符号秩检验应该被优先使用。

3. 与学生 t 检验的比较

如果总体分布是对称性的，且方差已知，那么符号检验、符号秩检验、学生 t 检验都可以被选择使用，因为在对称性分布情况下，均值与中位数相等。然而，学生 t 检验是建立在正态分布假设基础上，这是一个比对称性假设严格得多的假定条件。当样本数据的正态假定可靠，且数据是在定距尺度测量时，学生 t 检验作为一种最强有力的方法应该被优先选择。但当下列情况之一发生时，非参数检验的方法要优于参数检验方法。

(1) 样本数目很小；

(2) 唯一可以得到的数据是定类或定序尺度测量的；

(3) 作为一组数据的处理，样本的中位数似乎比均值更可靠；

(4) 对于所研究的问题来说，中位数是比均值更有代表性的位置参数；

(5) 总体很少或者几乎没有一个概率分布(对符号秩检验仅需要一个对称性假定)；

(6) 总体分布未知，但几乎很少类似于正态。

§2.5 游程检验

游程检验亦称连贯检验或串检验，是一种随机性检验方法，应用范围很广。例如生产过程是否需要调整，即不合格产品是否随机产生；奖券的购买是否随机；期货价格的变化是否随机等等。若事物的发生并非随机，即有某种规律，则往往可寻找规律，建立相应模型，进行分析，作出适宜的决策。

一、普通的游程检验(Ordinary Runs Test)

1.游程的含义

一个可以两分的总体，如按性别区分的人群，按产品是否有毛病区分的总体等等，随机从中抽取一个样本，样本也可以分为两类：类型Ⅰ和类型Ⅱ。若凡属类型Ⅰ的给以符号A，类型Ⅱ的给以符号B，则当样本按某种顺序排列(如按抽取时间先后排列)时，一个或者一个以上相同符号连续出现的段，就被称作游程，也就是说，游程是在一个两种类型的符号的有序排列中，相同符号连续出现的段。例如，将某售票处排队等候购票的人按性别区分，男以A表示，女以B表示。按到来的时间先后观察序列为：$AABABB$。在这个序列中，AA为一个游程，连续出现两个A；B是一个游程，领先它的是符号A，跟随它的也是符号A；显然，A也是一个游程，BB也是一个游程。于是，在这个序列中，A的游程有2个，B的游程也有2个，序列共有4个游程。每一个游程所包含的符号的个数，称为游程的长度。如上面的序列中，有一个长度为2的A游程、一个长度为2的B游程，长度为1的A游程、B游程也各有1个。

2.基本方法

随机抽取的一个样本，其观察值按某种顺序排列，如果研究所关心的问题是：被有序排列的两种类型符号是否随机排列，则可以建立双侧备择，假设组为

H_0: 序列是随机的

H_1: 序列不是随机的

如果关心的是序列是否具有某种倾向，则应建立单侧备择，假设组为

H_0: 序列是随机的

H_+: 序列具有混合的倾向

或

H_0: 序列是随机的

H_+： 序列具有成群的倾向

为了对假没作出判定，被收集的样本数据仅需定类尺度测量，但要求进行有意义的排序，按一定次序排列的样本观察值能够被变换为两种类型的符号。如某售票处按到来的先后顺序排队购票的人，按性别分别记作 A、B 两种类型的符号，可以得到一个序列：$AABABB$。第一种类型的符号数目记作 m，第二种记作 n，则 $N = m + n$。

检验统计量。在 H_0 为真的情况下，两种类型符号出现的可能性相等，其在序列中是交互的。相对于一定的 m、n，序列游程的总数应在一个范围内。若游程的总数过少，表明某一游程的长度过长，意味着有较多的同一符号相连，序列存在成群的倾向；若游程总数过多，表明游程长度很短，意味着两个符号频繁交替，序列具有混合的倾向。因此，无论游程的总数过多或过少，都表明序列不是随机的。根据两种类型符号的变化，选择的检验统计量为 U，

U = 游程的总数目

确定 P 值。游程总数目 U 的抽样分布在附表 Ⅷ 中给出。序列中数目比较少的符号记作类型 Ⅰ，数目多的符号为类型 Ⅱ。对于 $m \leqslant n$，且 $m + n \leqslant 20$，或 $m \leqslant n \leqslant 12$ 时，可以在附表 Ⅷ 中查找到相应的 P 值。若 P 相对于给定的显著性水平 α 很小，则数据不支持 H_0；若足够大，则不拒绝 H_0；表 2—27 是判定的指导表。

表 2—27 **游程检验判定指导表**

备择假设	P 值 （附表 Ⅷ）
H_+： 序列具有混合的倾向	U 的右尾概率
H_-： 序列具有成群的倾向	U 的左尾概率
$H_\pm$： 序列是非随机	U 的较小尾巴概率的 2 倍

当 $m + n = N > 20$ 或 $m > 12, n > 12$ 时，检验统计量 U 近似均值为 $1 + 2mn/N$，标准差为 $\sqrt{2mn(2mn - N)/N^2(N-1)}$，正态

分布通过连续性修正，计算 Z_L 或 Z_R，查附表Ⅳ，可以得到相应的 P 值。Z_L、Z_R 计算如(2.12)式。表 2—28 是判定指导表。

$$Z_L = \frac{U + 0.5 - 1 - 2mn/N}{\sqrt{2mn(2mn - N)/N^2(N-1)}}$$
$$Z_R = \frac{U + 0.5 - 1 - 2mn/N}{\sqrt{2mn(2mn - N)/N^2(N-1)}} \quad (2.12)$$

表 2—28　　游程检验判定指导表

	备择假设	P 值　(附表Ⅳ)
H_+:	序列具有混合的倾向	Z_R 的右尾概率
H_-:	序列具有成群的倾向	Z_L 的左尾概率
H_+:	序列是非随机	Z 的右尾概率的 2 倍

表 2—28 中，Z 的取值如下：

$$Z = \begin{cases} -Z_L & 若\quad U < 1 + 2mn/N \\ Z_R & 若\quad U > 1 + 2mn/N \end{cases}$$

3. 应用

【例 2.19】　某旅游点该年气温偏差是否随机

某旅游点该年二月份的气温，连续 10 天被记录，每天的最高气温与历史上同期最高气温平均值比较，高于均值记作 A，低于均值记录作 B，结果 10 天的气温依次记录为 $AABABBAAAB$。使用 $\alpha = 0.05$ 的水平，检验高温的偏差是否随机。

分析：根据检验的要求，建立的假设组为

H_0:　序列是随机的

H_1:　序列非随机

在调查结果按时间排序的两符号序列中，有 6 个 A，4 个 B，因为符号 B 的数目少，因而 B 为类型Ⅰ，即有 $m = 4$，$n = 6$。序列 A 的游程个数为 3，B 游程个数也是 3，游程总数 $U = 6$。查附表Ⅶ，$m = 4$，$n = 6$，$U = 5$ 时，$P = 0.405$；$m = 4$，$n = 6$，$U = 7$ 时，$P = 0.310$。显然，$m = 4$，$n = 6$，$U = 6$ 时，P 值在 0.310 与 0.405 之间，这对于

$n = N = 30 > 20$，因此采用正态近似。根据表 2—28 的提示，合适的 P 值是相对于 Z_L 的左尾概率。由于 $U = 4$，所以 Z_L 可以计算得到

$$Z_L = \frac{4 + 0.5 - 1 - 2 \times 7 \times 23/30}{\sqrt{2 \times 7 \times 23(2 \times 7 \times 23 - 30)/30^2 \times 29}}$$
$$= -217/57 = -3.81$$

在附表 N 中，$Z_L = -3.81$ 的左尾概率小于 0.0002，这相对于显著性水平 $\alpha = 0.05$ 来说是一个极小的值，因而数据不支持 H_0。根据这次调查的结果，产品的缺陷有成群产生的倾向，因此，应每天频繁抽取小样本检验，以保证估计的可靠。

二、基于上、下游程的检验

上、下游程(Runs Up and Down) 亦称升降串。这个检验不是单纯地用两个符号的多少以及游程的数目来进行，而是利用每个观察值与紧挨其前面的一个数值比较大小，决定升、降，利用形成的升降串进行检验。因此它比普通的游程检验能够提供更多的信息。

1. 上、下游程的含义

随机抽样得到的观察值按某一顺序排列，序列中每个观察值与其前面的一个数值比较，如果前面的数值较小，就构成一个上升串，即一个上游程；如果前面的数值比较大，就构成一个下降串，即一个下游程。一个上游程中包含的观察值数目，就是游程的长度。例如，某一序列的观察值为 7，15，1，2，5，8。这个序列有一个长度是 1 的上游程，因为第二个值 15 比前面的值 7 要大，而比后面的值 1 也要大，紧跟着是一个长度为 1 的下游程，然后是一个长度为 3 的上游程。如果用 +、— 表示上升或下降的变化方向，那么上面序列的变动结果是 +，—，+，+，+。这个序列观察值的数目，即样本数据的个数 $N = 6$，上、下游程的总数为 3。

2. 基本方法

和普通游程检验类似，如果研究的问题是序列是否随机，可以建立双侧备择，若关心的是序列是否有某种倾向，则建立单侧备择。基于上、下游程的检验所建立的假设组为

H_0： 序列是随机的

H_1： 序列是非随机的

H_0： 序列是随机的 H_0： 序列是随机的

H_+： 序列具有混合倾向 H_-： 序列具有成群倾向

为对假设作出判定，所需要的数据至少是定序尺度测量的。在 H_0 为真的情况下，相继的观察值之间差值的符号为 + 或为 − 的可能性相等。因此，上、下游程的总数可以反映序列的变动。如果相同的符号成群，游程的总数就会太少，表明序列是有一个恒定方向的倾向，或顺序增加，或顺序减少；如果符号不断地变化，游程的总数就会很多，表明序列有经常的波动或循环移动或漂浮不定的变动。

检验统计量。N 个不同观察值按某一顺序排列后，可以得到一个 $N-1$ 个正号或负号组成的相应次序的序列，这个序列的上、下游程总数就是检验统计量，记作 V，即有

V = 上、下游程的总数

确定 P 值，检验统计量 V 与 U 的分布不同，V 的抽样分布在附表 Ⅸ 中给出。V 的取值范围从 1 到 $N-1$，N 是观察值即样本数据的数目。当 $N \leqslant 25$ 时，根据 N、V 可以在附表 Ⅸ 中查到相应的 P 值。表 2—29 是检验判定的指导表。

表 2—29 上、下游程检验判定指导表

备择假设	P 值 （附表 Ⅸ）
H_-： 序列有混合倾向	V 的右尾概率
H_-： 序列有成群倾向	V 的左尾概率
H_1： 序列是非随机的	V 的较小尾巴概率的 2 倍

当 $N > 25$ 时，检验统计量 V 近似正态分布，均值为 $(2N-1)/3$，标准差是 $\sqrt{(16N-29)/90}$。通过连续性修正，计算得到 Z_L，Z_R，查

排列，中位数是25，则高于中位数的编号为1，低于中位数的编号为0。将表2—31中的按观察顺序记录的分数换成相应的符号0、1，可以得到$m=12,n=12,N=m+n=24,U=10$。查附表Ⅷ，得到$P=0.150$。由于是双侧检验，因而P应取$2(0.0150)=0.30$，对于显著性水平$\alpha=0.05$，P已足够大，表明数据支持H_0，即调查结果没有受儿童之间相互通气的影响，是随机的。

表2—31 按观察顺序记录的儿童好斗程度评分

编号	分数	编号	分数	编号	分数	编号	分数
1	31	7	12	13	15	19	86
2	23	8	26	14	13	20	61
3	36	9	43	15	78	21	13
4	43	10	75	16	24	22	7
5	41	11	2	17	13	23	6
6	44	12	3	18	27	24	8

上、下游程检验。将表2—31的调查评分，按升、降给以符号+、－，得到下面23个符号

$$-,+,+,-,+,-,+,+,+,-,+,+,$$
$$-,+,-,-,+,-,-,-,-,-,+$$

这一序列，包含上、下游程数目总共14个，即$V=14$。查附表Ⅸ，$N=24,V=14,P=0.2768$，由于是双侧检验，P应是$2(0.2768)=0.5536$。显然，对于显著性水平$\alpha=0.05$，P已足够大。因此，调查数据在5%的显著水平上不能拒绝H_0，表明调查结果没有受儿童相互间通气的影响。与普通游程检验的结论一致。

第三章　两个相关样本的非参数检验

某种统计检验方法应用时，不仅与数据的测量层次有关，还与抽样的特点有关。在抽取样本时有两种形式：相关的和独立的。若第一次抽样的所有样本某一属性的测量结果，不影响第二次抽样的所有样本同一属性的测量结果，则这种抽样是独立的；若一次抽样的测量结果影响另一次抽样测量结果，则这种抽样是相关的。本章介绍两个相关样本的非参数检验方法。

实际问题中，常会有这样的情况：某种药物对治疗某种疾病是否有效，某种训练方法是否能提高人的能力，某次宣传是否能提高人们的认识，价格上涨是否会产生不良影响等等。当研究者希望知道两种处理结果是否相同，或哪种更好时，往往需要采用两个样本的统计检验。这时，常用经过处理的一组和未经处理的一组比较，或者一种处理组与另一种处理组相比较，为了避免或尽量减少由于其它因素影引起的两组之间的附加差异，研究中通常采用两个相关样本。例如，在接受两种不同训练方法的人员中，由于智力、接受能力、耐力等方面的不同，会导致不同处理的结果产生差异，这不是所要研究的问题，而是其它因素影响产生的附加差异。这些因素在实施不同处理前必须排除。为获取相关样本，常应用两种方式。一是让每一研究对象作为自身的对照者，一是将研究对象两两配对，分别给每一对两个成员以不同处理。在进行配对时，应让每一对在可能影响处理结果的其它因素方面尽量相似，以尽量避免和减小附加差异。一般来说，用研究对象自身作为对照者要优于配对方法。因为很难在配对过程中，完全控制住其它的影响因素。

§ 3.1　符号检验

一、基本方法

设有两个连续总体 X、Y，累积的分布函数分别为 $F(x)$，$F(y)$。随机地分别从两个总体中抽取数目为 n 的样本数据 $(x_1,x_2,\cdots,x_n)$ 和 $(y_1,y_2,\cdots,y_n)$，将它们配对得到 (x_1,y_1)，(x_2,y_2)，$\cdots$，(x_n,y_n)。若研究的问题是它们是否具有相同的分布，即 $F(x)=F(y)$ 是否成立。由于 X、Y 的总体分布未知，而研究也并不关心它们的具体分布形式，只是关心分布是否相同。因而，可以采用位置参数进行判断。若两个样本的总体具有相同分布，则中位数应相同，即在 n 个数对中，x_i 大于 y_i 的个数与 x_i 小于 y_i 的个数应相差不多。若 P 表示概率，则建立的假设组为

$$H_0:\quad P(x_i>y_i)=P(x_i<y_i)\qquad \text{对所有 } i$$
$$H_1:\quad P(x_i>y_i)\neq P(x_i<y_i)\qquad \text{对某一 } i$$

如果关心的是某一总体中位数是否大于另一总体中位数，则可建立单侧备择，假设组为

H_0:　$P(x_i>y_i)=P(x_i<y_i)$　H_0:　$P(x_i>y_i)=P(x_i<y_i)$

H_+:　$P(x_i>y_i)>P(x_i<y_i)$　H_-:　$P(x_i>y_i)<P(x_i<y_i)$

在 H_+ 下，x_i 有大于 y_i 的趋向，在 H_- 下，y_i 有大于 x_i 的趋向。

为对假设作出判定，所需的数据至少是定序尺度测量。与单样本的符号检验相同，两个相关样本的符号检验也定义 S_+、S_- 为检验统计量。S_+ 为 x_i、y_i 差值符号是正的数目，S_- 为差值符号是负的数目，$S_++S_-=n$。若 H_0 为真，$x_i>y_i$ 的配对数目与 $x_i<y_i$ 的配对数目相等，也就是 S_+ 与 S_- 的数值相等。由于 S_+、S_- 的抽样分布是二项分布 $B(n,\frac{1}{2})$，n 是配对数目，$\frac{1}{2}$ 是各自出现的概率，因而合适的 P 值能够在附表 Ⅵ 中查找到。若 P 值相对于显著性水平 α 很小，则数据不支持 H_0。判定指导表可参见表 2—12。当 n

$\leqslant 20$ 时，查找附表 Ⅵ，而当 $n > 20$ 时，应采用(2.7)式的计算公式，得到 $Z_{-,R}$，$\dot{Z}_{-,R}$，利用正态近似，查附表 Ⅳ，找到合适的 P 值。

二、应用

【例 3.1】 教学参考资料对于指导学生自学是否有效

为帮助学生通过自学提高对知识的掌握，编辑了符合教学大纲的教学参考资料。针对某一概念的掌握进行实验，随机选取了 15 名学生，他们在使用参考资料之前的得分如表 3—1。学习参考资料后，重新对这一概念进行测试，得分也列在表 3—1 中。检验这部参考资料是否促进学生掌握知识。

表 3—1 学生两次测试成绩

学生编号	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
第一次成绩	2	2	2	2	3	3	3	3	3	3	2	3	3	2	3
第二次成绩	3	4	2	3	2	3	4	4	2	4	4	3	4	4	4

分析：由于关心的是学生使用参考资料后是否对概念的掌握更好了，因此应建立单侧备择，假设组为

$$H_0: \quad P(x_i > y_i) = P(x_i < y_i)$$

$$H_-: \quad P(x_i > y_i) < P(x_i < y_i)$$

x_i 为第一次检查的成绩，y_i 为第二次测试的成绩。由表 3—1 的分数可知，$S_+ = 2$，$S_- = 10$，有 3 对差值为 0，无法记 + 或 − 号，因此，符号总数目 $n = S_+ + S_- = 12$，而不是 15。在附表 Ⅵ 中查找，$n = 12$，$S_- = 10$ 的右尾概率 $P = 0.0193$，对于显著性水平 $\alpha = 0.05$，显然 P 值够小，因而数据不支持 H_0，即学生通过自学参考资料对提高知识有益，表明教学参考资料对促进学生掌握知识是有效的。

【例 3.2】 母亲是否比父亲更强烈地认识到父亲教育孩子的重要

在研究父亲对儿童发展的影响作用时，进行了一项调查。随

机抽选 17 对夫妇，其家庭均因两地分居，在孩子出生时，由母亲一人抚养，一年后团聚。针对父亲的教育对孩子影响较大这一问题询问，结果如表 3—2。表中是对父、母态度的评分，5 分代表最强烈的认识，1 分相对最弱，x_i 是父亲的态度，y_i 是母亲的态度。

表 3—2　**父母对父亲教育孩子重要的认识**

编号	x_i	y_i	编号	x_i	y_i	编号	x_i	y_i
1	2	4	7	3	5	13	5	4
2	3	4	8	3	3	14	2	5
3	3	5	9	2	1	15	5	5
4	3	5	10	3	5	16	3	5
5	3	3	11	2	5	17	1	5
6	3	2	12	2	5			

分析：这是一个配对样本的例子，例 3.1 是将研究对象自身作为对照者，形成两个相关样本。此例中是采用夫妇配对，讨论同一孩子和同一家庭的情况，因而配对有意义，符合符号检验的条件。由于关心的是母亲比父亲更强烈地认识到父亲教育孩子重要，因此是单侧检验，建立的假设组为

$$H_0:\quad P(x_i > y_i) = P(x_i < y_i)$$

$$H_-:\quad P(x_i > y_i) < P(x_i < y_i)$$

因为预言母亲比父亲的感受更强烈，因而评分应是母亲比父亲更高，故 $x_i - y_i < 0$ 的数目应较多，备择假设选择了 H_-。由表 3—2 可知，其中 3 对夫妇的态度一样，分数相等，差值为 0，不记符号，因此，$n = 14, S_+ = 3, S_- = 11$。查附表 Ⅵ，$n = 14, S_- = 11$ 右尾概率 $P = 0.0287$。对于显著性水平 $\alpha = 0.05$，P 值足够小。因此，数据不支持 H_0，调查结果表明，母亲确实比父亲更强烈地感到父亲教育孩子很重要。

配对样本位置的符号检验，在第二章中已介绍，实际上那也是两个相关样本的情况。

§3.2 Wilcoxon 符号秩检验

一、基本方法

两个相关样本的 Wilcoxon 符号秩检验也是用来检验配对样本是否有差异的方法。它不仅借助于两个样本差值的符号，而且利用了差值的大小，因此，它比符号检验有更精确的判断。

设 X、Y 是两个连续总体，且均具有对称的分布，随机地分别从两个总体中抽取 n 个观察值，组成 n 个数对 (x_1, y_1)，(x_2, y_2)，…，(x_n, y_n)。记 $D_i = x_i - y_i$，若 X 与 Y 具有相同的分布，则等式

$$P(D_i > 0) = P(D_i < 0)$$

成立，即 x_i 大于 y_i 的概率与 x_i 小于 y_i 的概率相等。这也意味着全部差值 D_i 的中位数等于零。因此，零假设也可以是

$$H_0: \quad D_i \text{ 的中位数} = 0$$

这与 §2.4 中配对样本的符号秩检验是一致的，当研究的问题仅关心两个总体的分布是否相同，或说两个总体中位数是否相同时，采用双侧备择；若 X、Y 之间的相互关系中，存在某种趋势，则应建立单侧备择。如果认为 x_i 的大多数值大于相应的值 y_i，那么单侧备择为

$$H_+: \quad P(D_i > 0) > P(D_i < 0)$$

或

$$H_+: \quad D_i \text{ 的中位数} > 0$$

如果认为 y_i 的值大多数大于相应的 x_i 的值，则单侧备择为

$$H_-: \quad P(D_i > 0) < P(D_i < 0)$$

或

$$H_-: \quad D_i \text{ 的中位数} < 0$$

若将差值 D_i 的总体中位数记作 M_D，那么，两个相关样本的 Wilcoxon 符号秩检验与 §2.4 中配对样本位置的符号秩检验基本

方法相同，判定假设是否成立的原则也一样，确定P值也可以按表2—22，表2—23的原则去做。

二、应用

【例3.3】 幼儿园的生活对孩子的社会知识是否有影响

有人认为儿童上幼儿园有助于其认识社会，有人则认为儿童在家一样可以获得社会知识。为了解它们是否存在差异，对8个同性孪生儿童进行实验，随机指定8对中一个上幼儿园，另一个则在家。经过一个时期后，通过对他们询问，给他们分别作出评价。评分结果如表3—3。

表3—3　　8对儿童的社会知识成绩

配对号	上幼儿园儿童	在家儿童	配对号	上幼儿园儿童	在家儿童
1	78	62	5	76	80
2	70	58	6	72	73
3	67	63	7	85	82
4	81	77	8	83	78

分析：虽然可以相信得分多的孩子比得分少的孩子社会知识要多，但由于是定距尺度测量，无绝对零值，因此不能认为得80分的孩子社会知识是得40分的孩子的2倍，也不能认为80分与60分的社会知识之差一定是60分与50分之差的2倍。但是，可以肯定，80分与60分所代表的社会知识之差一定大于60分和50分之间的差。所以将分数差值的绝对值排序是有意义的，这样就有能够运用Wilcoxon符号秩检验判定在家和上幼儿园对孩子的社会知识是否有影响。由于只关心两组儿童的社会知识是否有差异。因此应建立双侧备择。假设组为

H_0:　D_i的中位数 $= 0$

H_1:　D_i的中位数 $\neq 0$

用文字表述为

H_0:　在家和上幼儿园的儿童社会知识没有差异

H_1:　两组儿童的社会知识有差异

根据表 3—3 的数据计算 $|D|$ 及 T_+、T_-，计算过程如表 3—4。

表 3-4　　检验统计量计算表

x	y	$D=x-y$	$\|D\|$	$\|D\|$ 的秩	D 的符号
78	62	4	4	5	+
70	58	12	12	8	+
67	63	44	5	5	+
81	77	4	4	5	+
76	78	−2	2	2	−
72	73	−1	1	1	−
85	82	3	3	3	+
83	78	5	5	7	+

由表可知

$$T_+ = 5+8+5+5+3+7=33$$

$$T_- = 3$$

根据 $n=8$，T_+ 与 T_- 中较大者 $T_+=33$，在附表 Ⅶ 中查找相应 P 值，得到 $T_+=33$ 右尾概率 $P=0.020$，这一概率的 2 倍为 0.04，对于显著性水平 $\alpha=0.05$，P 值显然较小，故调查结果不支持 H_0。这表明在 5% 的显著性水平上，拒绝零假设，在家和上幼儿园对儿童的社会知识有影响。从计算结果看出，+ 号大大多于 − 号，表明上幼儿园的儿童社会知识成绩普遍高于在家的儿童。

第四章　两个独立样本的非参数检验

利用两个相关样本进行研究，对某些问题是很方便的。但现实中要做到很好配对并不容易。若由于配对不当或无法配对，第三章介绍的方法不能使用。这时可以运用两个独立样本的非参数检验方法。两个独立样本可以各自从两个总体中随机抽选获得，也可以对随机抽样的一个样本诸元素，随机分别实施两种处理而形成。两个样本的观察值数目不一定要求相同。

分析两个独立样本常用的参数方法是 t 检验，即对两样本均值是否相等的检验。t 检验要求分析的数据是来自方差相等的正态分布总体的独立观察结果，并至少是定距尺度测量。实际研究中，由于种种原因，往往不符合 t 检验的条件或并不在乎是否符合条件，t 检验无法使用。当研究所处理数据的测量低于定距尺度，或不愿做严格的假设而使结论更具普遍性或无从得知总体分布，或总体分布非正态等等，选用非参数检验方法往往更为有效。

§4.1　Mann-Whitney-Wilcoxon 检验

Mann-Whitney-Wilcoxon 检验，常译为曼-惠特尼-维尔科克森检验，简写为 M-W-W 检验，亦称 Mann-Whitney U 检验

一、基本方法

两个变量 X、Y，其累积分布函数分为 F_x 和 F_y，若考察两个总体 F_x 与 F_y 是否有差异，可建立零假设

$$H_0:\quad F_x(u) = F_y(u) \qquad \text{对所有的 } u$$

如果考察的是总体在位置上是否不同，也可以采用上述零假设。在实际问题中，应用 U 检验常是考察两个总体的中心是否相同。若 M_x、M_y 分别是 X、Y 总体的中位数，则零假设可为

$$H_0:\quad M_x = M_y$$

若当 X、Y 都存在算术平均数时，也可以建立两个均值相等的零假设。当研究只关心两个总体中位数是否有差异时，采用双侧备择；若认为 X 的值可能大于 Y 的值这种趋势或倾向存在时，应建立单侧备择

$$H_+: \quad M_x > M_y$$

若相反，X 的值可能平均地小于 Y 的值，则单侧备择为

$$H_-: \quad M_x < M_y$$

U 检验可建立的假设组为

$$H_0: \quad M_x = M_y$$

$$H_1: \quad M_x \neq M_y$$

$$H_0: \quad M_x = M_y \qquad H_0: \quad M_x = M_y$$

$$H_+: \quad M_x > M_y \qquad H_+: \quad M_x > M_y$$

为了对假设作出判定，分析的数据应是两个相互独立的随机样本：$x_1, x_2, \cdots, x_m$ 和 $y_1, y_2, \cdots, y_n$。它们分别从连续总体 F_x 和 F_y 中随机抽取出来。数据的测量层次至少是定距尺度，若是定序尺度测量，则每个观察值的相对大小应能被确定。

如果 H_0 为真，那么将 m 个 X、n 个 Y 的数据，按数值的相对大小从小到大排序，X、Y 的值应该期望被很好地混合，这 $m+n=N$ 个观察值能够被看作来自于共同总体的一个单一的随机样本。若大部分的 Y 大于 X，或大部分的 X 大于 Y，将不能证实这个有序的序列是一个随机的混合，将拒绝 X、Y 来自一个相同总体的零假设。在 X、Y 混合排列的序列中，X 占有的位置是相对于 Y 的相对位置，因此等级或称秩是表示位置的一个极为方便的方法。在 X、Y 的混合排列中，等级 1 是最小的观察值，等级 N 是最大的。若 X 的等级大部分大于 Y 的等级，那么数据将支持 H_+；而 X 的等级大部分小于 Y 的等级，则数据将支持 H_-。无论上面哪一种情况发生，双侧的备择 H_1 都将被支持。

检验统计量。根据上面的基本原理，U 检验定义的检验统计量为

$T_x = X$ 等级的和，即 X 的秩和

$T_y = Y$ 等级的和，即 Y 的秩和

由于 X、Y 混合序列的等级和为

$$1 + 2 + \cdots + N = N(N + 1)/2$$

所以，$T_x + T_y = N(N + 1)/2$。从而可得

$$T_x = N(N + 1)/2 - T_y$$

M-W-W 检验可以直接用 T_x 作为检验统计量，也可以用 U。U 被定义为

$$U = T_x - m(m + 1)/2$$

例如，一种药物有效性的实验，一组为实验组，一组为对照组。实验结果评分为

实验组(X)	8	12	18	
对照组(Y)	6	9	11	13

将这些评分按增加的次序排列，并注意哪一个评分为 X 组的，哪一个为 Y 的，同时给出秩，结果为

	6	8	9	11	12	13	18
秩	1	2	3	4	5	6	7
组别	Y	X	Y	Y	X	Y	X

由上面排序可知

$$T_x = 2 + 5 + 7 = 14$$

$$T_y = 1 + 3 + 4 + 6 = 14$$

因为，$m = 3$，$n = 4$，故 $N = m + n = 7$，$T_x = N(N + 1)/2 - T_y = 28 - 14 = 14$，$U = Tx - m(m + 1)/2 = 14 - 6 = 8$。$U$ 是 Y 的评分领先于 X 的总次数。在上面的例子中，先于 X 的评分 8 的 Y 的评分有 1 个，先于 X 的评分 12 的 Y 的评分值有 3 个：6、9、11，而先于最后一个 X 的评分值的 Y 值有 4 个：6、9、11 和 13。在 X、Y 混合的序列中，Y 的评分值先于 X 的总次数为 $1 + 3 + 4 = 8$，这就是检验统计量 U。一般情况下，当两组样本数据数目不等时，较少数目的组记为 X，即 $m \leqslant n$。

确定 P 值。在 M-W-W 检验中的统计量 T_x，当 $m \leqslant n$ 时，取值为整数，范围为 $m(m+1)/2$ 到 $(2N-m+1)/2$。这时，T_x 的抽样分布关于其均值 $m(N+1)/2$ 是对称的。在 $m \leqslant n \leqslant 10$ 的情况下，可以根据 m、n 以及 T_x 的值查附表 X，得到相应的 P 值。当 $T_x \leqslant m(N+1)/2$ 时，查左尾概率，若 $T_x \geqslant m(N+1)/2$ 则查右尾概率。表 4—1 是检验的判定指导表。当 m,n 均大于 10 时，T_x 近似于均值为 $m(N+1)/2$，标准差为 $\sqrt{mn(N+1)/12}$ 的正态分布。这时，通过连续性校正，利用(4.1)式计算得到 $Z_{x,L}$、$Z_{x,R}$，查附表 Ⅳ，得到相应的 P 值．判定指导表见表 4—2。

表 4-1　　　　U 检验判定指导表

备择假设	P 值(附表 X)
H_+: $M_x > M_y$	T_x 的右尾概率
H_-: $M_x < M_y$	T_x 的左尾概率
H_1: $M_x \neq M_y$	T_x 较小概率的 2 倍

$$
\begin{aligned}
Z_{x,L} &= \frac{T_x + 0.5 - m(N+1)/2}{\sqrt{mn(N+1)/12}} \\
Z_{x,R} &= \frac{T_x - 0.5 - m(N+1)/2}{\sqrt{mn(N+1)/12}}
\end{aligned}
\qquad (4.1)
$$

表 4-2　　　　U 检验判定指导表

备择假设	P 值　(附表 Ⅳ)
H_+: $M_x \gtrsim M_y$	$Z_{x,y}$ 的右尾概率
H_-: $M_x < M_y$	$Z_{x,L}$ 的左尾概率
H_1: $M_x \neq M_y$	Z 的右尾概率的 2 倍

表 4—2 中 Z 的取值，对不同 T_x 有所不同，定义如下

$$
Z = \begin{cases} -Z_{x,L} & T_x < m(N+1)/2 \\ Z_{x,R} & T_x > m(N+1)/2 \end{cases}
$$

二、应用

【例 4.1】 某种药物对治疗肿瘤是否有效

选择 9 只白鼠，作为抗癌药物筛选的对象。9 只白鼠的基本条件相同，同时注射致癌物。然后随机选取其中 3 只进行抗癌药物处理。肿瘤的重量是检验药物有效性的一个指标。经过一个固定的时间周期后，将 9 只白鼠的肿瘤割除称重，结果如下（重量单位是克）

处理组(X)　0.94，1.56，1.15

控制组(Y)　1.20，1.63，2.26，1.87，2.20，1.30

分析：若该种抗癌药物有效，处理组白鼠肿瘤的重量应该小于控制组的平均重量。由于这些实验采用的是小样本，且为两个独立样本，数据测量为定比尺度，可运用 Mann-Whitney-Wilcoxon 检验。建立的假设组为

$$H_0:\quad M_x = M_y$$

$$H_-:\quad M_x < M_y$$

将肿瘤重量从小到大排序为

	0.94	1.15	1.20	1.30	1.56	1.63	1.87	2.20	2.26
秩	1	2	3	4	5	6	7	8	9
组	X	X	Y	Y	X	Y	Y	Y	Y

$T_x = 1 + 2 + 5 = 8, T_y = 3 + 4 + 6 + 7 + 8 + 9 = 37, m = 3, n = 6$。查附表 X，$m = 3, n = 6, T_x = 8$ 的左尾概率 $P = 0.048$，对于显著性水平 $\alpha = 0.05$，P 还不够大。因此，实验数据不支持 H_0，即 9 只白鼠的药物实验结果表明，在 5% 的显著性水平上，该抗癌药物对控制肿瘤有效。

三、同分的处理

在 $x_1, x_2, \cdots, x_m$ 与 $y_1, y_2, \cdots, y_n$ 排序时，若其均能精确测量，数据完全相同的可能性极小，但有时测量达不到很高的精度，因而

会出现相同的观察值。序列中观察值相同称作同分。观察值同分时，其秩为所占位置顺序号的算术平均数。若同分出现在一个样本之内，检验的精确性将不会受到影响。但若同分出现在两个样本之中，给这些同分值以相同的秩，将会降低检验的精确性。由于小样本情况下，这种降低不大，故通常忽略。但大样本时，应采用(4.2)式来校正 T_x 抽样分布的标准差，即当 m、n 均大于10时，无论在一个样本内出现同分，或两个样本间出现同分，或一个样本内与两个样本间均存在同分，应以(4.2)式替代(4.1)式中的分母。(4.2)式中的 u 是同分的观察值数目。如有两个样本的观察值分别是，X：1,2,5,5；Y：1,3,6,8,8,8 观察值中有2个1,2个5,3个8于是，$u_1 = 2, u_2 = 2, u_3 = 3$。

$$
\begin{aligned}
&\sqrt{\frac{mn(N+1)}{12} - \frac{mn(\Sigma u^3 - \Sigma u)}{12N(N-1)}} \\
&= \sqrt{\frac{mn}{12N(N-1)}[N(N^2-1) - (\Sigma u^3 - \Sigma u)]}
\end{aligned} \qquad (4.2)
$$

【例4.2】 问题按难易次序提问是否影响学生正确回答的能力

从心理学的角度看，按问题的难易程度顺序提问会影响学生正确回答的能力，从而影响他们的总分数。为检验这种观点，随机地将一班学生20人分成两组，每组10人。设计一组问题，分成 A、B 卷。A 卷是问题按从易到难的次序安排，B 卷相反，从最难到最易。两组学生分别回答 A、B 卷，考试被控制在完全相同的条件下进行，评分结果如下：

A：83，82，84，96，90，64，91，71，75，72

B：42，61，52，78，69，81，75，78，78，65

分析：这一问题可以考虑按考试分数的中位数来研究。若两组成绩的中位数相等，提问的次序对学生的成绩无影响，若中位数不相等则不敢认为没有影响。由于是小样本，并且为两个独立样本，因而可以运用M-W-W检验。这是一个单侧检验，单侧备择应

Y 总体的中位数之间差值不是零，而是某个特定的值 θ_0。当其它条件都与前面所述的相同时，应用 Mann-Whitney-Wilcoxon 检验的假设可以写为

$$H_0: \quad M_x - M_y = \theta_0$$
$$H_1: \quad M_x - M_y \neq \theta_0$$

单侧检验为

$$H_0: \quad M_x - M_y = \theta_0 \qquad H_0: \quad M_x - M_y = \theta_0$$
$$H_+: \quad M_x - M_y > \theta_0 \qquad H_-: \quad M_x - M_y < \theta_0$$

上述的零假设也可以写成

$$H_0: \quad (M_x - \theta_0) - M_y = 0$$

如果 $X - \theta_0$ 的总体分布与 Y 的总体分布相同，即有

$$F_y(u) = F_x(u - \theta_0) \qquad \text{对所有的 } u$$

那么，M_x 为 X 的中位数，M_y 为 Y 的中位数，θ_0 就是 M_x 与 M_y 的差值。定义 $x' = x - \theta_0$，即对 x 的 m 个观察值有 $x'_1 = x_1 - \theta_0, x'_2 = x_2 - \theta_0, \cdots, x'_m = x_m - \theta_0$，那么，将 m 个 x' 的值与 n 个 Y 的值排序，分别给以相应的秩，可以应用 Mann-Whitney-Wilcoxon 检验。检验统计量

$$T_x' = x' \text{ 的秩和}$$

除了作这一变换外，其它检验步骤，原则完全同前面所述的一样。

§4.2 Wald-Wolfowitz 游程检验

Mann-Whitney-Wilcoxon 检验主要应用于检验两个样本是否来自具有相同位置的总体，是对两个总体在集中趋势方面有无差异的一种考察，而不研究其它类型方面的差异。Wald-Wolfowitz 游程检验则可以考察任何一种差异。Wald-Wolfowitz Runs Test 常译为沃尔德—沃尔福威茨连串检验或游程检验，简写为 W-W 串检验。

一、基本方法

设有 X、Y 的两个总体具有连续分布，其累积分布函数分别为

F_x,F_y。若考虑两个总体是否存在某种差异,即检验两个总体分布相同的零假设是否成立。建立的假设组为

$$H_0:\quad F_y(u)=F_x(u)\qquad 对所有的\ u$$

$$H_1:\quad F_y(u)\neq F_x(u)\qquad 对某个\ u$$

为对假设作出判定,需要从 X 中随机抽取 m 个数据 $x_1,x_2,\cdots,x_m$,从 Y 中随机抽取了 n 个数据 $y_1,y_2,\cdots,y_n$。数据的测量层次至少要是定序尺度。将两个独立样本的 $m+n=N$ 个数据按大小排列,即将所有 N 个数据排成一个有序的序列,确定这个序列的游程数,也就是连串数。一个游程定义为取自同一样本的一串相连的数据。例如,观察两组学生的考试成绩如下

X: 72, 78, 63

Y: 65, 79, 82, 85

将 7 个分数排列成一个从小到大的序列为

63	65	72	78	79	82	85
X	Y	X	X	Y	Y	Y

观察 X、Y 出现的次序以确定游程数。序列中有 4 个游程;一个由来自 X 的 63 分构成的游程,随后是一个由来自 Y 的 65 分构成的游程,再后是由来自 X 的两个分数 78 和 72 构成的游程,最后是三个来自 Y 的分数构成的 1 个游程。如果 H_0 为真,则两个样本的数据期望能相互混合地排列,游程数会相对较大。若 X 的游程或 Y 的游程过长,也就是来自同一总体的数据在有序的序列中过多的相互连接,则游程数将会相当小,这样,数据将不支持 H_0。所以,可以用序列的游程数作为检验统计量。定义 U 为 Wald-Wolfowitz 检验的统计量

$$U=游程的总数目$$

确定 P 值。当 $m+n=N\leqslant 20$ 时,与单样本游程检验相同,在附表 Ⅷ 中,依据 m、n 及 U 查找相应的 P 值。由于 Wald-Wolfowitz 检验通常是双侧检验,所以按表 2—27 的判定指

导原则确定 P 时，应选双侧备择。若 $m+n=N>20$ 或 $m>12$，$n>12$，则 U 的抽样分布近似正态分布，计算 Z，依据表 2—28 的原则查找相应的 P 值。

二、应用

【例 4.3】 问题的提问顺序是否对学生正确回答的能力有影响

沿用例 4.2 的资料，考察问题的提问顺序是否对学生成绩产生影响。

分析：由于只考察问题从易到难排序和从难到易排序是否会影响学生的成绩，且相互独立，因此可以用 Wald-Wolfowitez 游程检验。假设组为

$$H_0:\quad F_y(u)=F_x(u)$$

$$H_1:\quad F_y(u)\neq F_x(u)$$

用文字表述为

H_0：　从易到难提问和从难到易提问，学生的成绩没有差异

H_1：　两种提问顺序会造成学生的成绩有差异

将实验数据即学生考试成绩从小到大排序得到

42	52	61	64	65	69	71	72	75	75
Y	Y	Y	X	Y	Y	X	X	X	Y
78	78	78	81	82	83	84	90	91	96
Y	Y	Y	Y	X	X	X	X	X	X

从上面结果可知，序列的游程总数目 $U=6$。在附表 Ⅷ 中，$m=10$，$n=10$，$U=6$ 的概率为 0.019。由于是双侧检验，相应的 P 值应是 $2\times0.019=0.038$。对于显著性水平 $\alpha=0.05$，显然 P 还不够大，因此，数据不支持 H_0，即提问的顺序对学生正确回答问题的能力有影响。

三、同分的处理

采用 Wald-Wolfwitz 游程检验与 Mann-Whitney U 检验一样，均假设总体是连续分布，因而若能精确测量，观察值不会有同分出现。但实际上，测量有时很难极准确，所以常会有同分出现。如果同分值来自同一个样本，游程数 U 不会受到影响，如例 4.3 中来自 Y 的 3 个 78 分，无论怎么排序都是构成 1 个游程。但同分值来自两个样本时，U 就可能会受到影响，并影响最后的结论。例 4.3 中的两个 75 分，分别来自 X、Y，在例中是先排的 X、再排 Y。若来自 Y 的排在先，来自 X 的排在后，序列的游程总数将不是 6，而是 $U=8$。在附表 Ⅷ 中，$m=10$，$n=10$，$U=8$，$P=0.128$。这种情况下，数据支持 H_0，得出与前面相反的结论。在运用 Wald-Wolfowitz 游程检验时，若同分值来自两个不同样本，一般应将各种排序的可能性都进行考察，分别计算每种情况下的游程总数 U，并查找相应的 P 值。如果得出的结论一致，表明同分没有带来什么问题；如果得到的结论不一致，可以将几个 P 值求简单平均数，以此作为是否拒绝 H_0 的依据。如例 4.3 中，一种排序的 $P=0.019$，另一种情况 $P=0.128$，则可取二者的简单算求平均数 $(0.128+0.019)/2=0.0735$ 作为相应的 P 值，决定是否支持出 H_0。显然，按照这个 P 值，在 0.05 的显著性水平上，数据支持 H_0。如果同分在两个样本之间多次出现，U 实际上是不确定的，因而不宜采用 Wald-Wolfowitz 游程检验。

§4.3 两样本的 χ^2 检验

单样本的 χ^2 检验方法可以推广到对两个独立样本的总体差异性的检验。

一、基本方法

分别从两个分布函数为 $F_1(x)$ 和 $F_2(x)$ 的总体中，随机抽取 n_1 和 n_2 个样本数据，利用样本值推断两个总体是否具有某种差

异，可以建立假设组

$$H_0:\quad F_1(x)=F_2(x)\qquad \text{对所有 } x$$

$$H_1:\quad F_1(x)\neq F_2(x)\qquad \text{对某个 } x$$

在具体研究某种特性的差异时，零假设和备择假设可以具体化。如检验不同性别儿童对电视台的几种少儿节目是否有偏好，检验不同文化程度的青年对职业的选择有否不同，检验不同性别的学生大学升学率是否有差异等等，可以用文字表述零假设和备择假设。

为了对假设作出判定，所需要的数据是两个样本，测量层次最低可为定类尺度。对每一个样本的数据都可以分为 r 个组，$r\geqslant 2$。若第一个样本数据各个组的观察频数分别记作 $f_{11},f_{21},\cdots,f_{r1}$；第二个样本各组的观察频数分别记作 $f_{12},f_{22},\cdots,f_{r2}$，那么，任一样本某组的观察频数可以用 f_{ij}，表示，其中 $i=1,2,\cdots,r$，$j=1,2$。第一个样本的观察值数目 $\Sigma f_1=n_1$，第二个样本观察值数目 $\Sigma f_2=n_2$，两个样本观察值总数目 $N=n_1+n_2$。两个总体中与两个样本观察频数相对应的期望频数分别记作 $e_{11},e_{21},\cdots,e_{r1}$ 和 $e_{12},e_{22},\cdots,e_{r2}$。第 i 组两个样本的观察频数和记作 $f_{i\cdot}$，那么

$$f_{i\cdot}=f_{i1}+f_{i2}$$

如果 H_0 为真，那么第一个样本第 i 组的期望频数

$$e_{i1}=n_1\frac{f_{i\cdot}}{N}\qquad (4.3)$$

第二个样本第 i 组的期望频数为

$$e_{i2}=n_2\frac{f_{i\cdot}}{N}=(N-n_1)\frac{f_{i\cdot}}{N}=\frac{f_{i\cdot}}{N}-e_{i1}$$

于是，两个样本的各组观察频数与相应的期望频数可以归纳如表4—4。若 H_0 为真，观察频数 f_{ij} 与期望频数 e_{ij} 应相等。因此，f_{ij} 与 e_{ij} 越接近，即其差值越小，表明 H_0 为真的可能性越大。对于两个独立样本，可以用 $(f_{i1}-e_{i1})^2/e_{i1}$ 和 $(f_{i2}-e_{i2})^2/e_{i2}$ 之和的大小来判定是否拒绝 H_0。

检验统计量。两个独立样本的 χ^2 拟合优度检验的检验统计量定义为 Q。它由下式计算

$$Q = \sum_{i=1}^{r} \frac{(f_{i1} - e_{i1})^2}{e_{i1}} + \sum_{i=1}^{r} \frac{(f_{i2} - e_{i2})^2}{e_{i2}} \qquad (4.4)$$

(4.4) 式也可以写成

$$Q = \sum_{i=1}^{r} \sum_{j=1}^{2} \frac{(f_{ij} - e_{ij})^2}{e_{ij}} \qquad (4.5)$$

确定 P 值。Q 统计量近似地遵从自由度 $df = r - 1$ 的 χ^2 分布。在附表 I 中，根据自由度和给定的显著性水平 α，查找相应的临界值 $\chi^2_\alpha(r-1)$。若 $Q \geqslant \chi^2_\alpha(r-1)$，则 Q 统计量落入拒绝域，即在显著性水平 α 上拒绝 H_0。反之，$Q < \chi^2_\alpha(r-1)$，则不能拒绝 H_0。对假设的判定还可以这样进行。在附表 I 中，根据 df、Q，查到 H_0 成立时近似的 P 值。若 P 对于给定的显著性水平 α 足够大，则支持 H_0，表明事件发生的概率很大，反之，P 值较小，则拒绝 H_0。

表 4—4　　　　χ^2 检验频数表

组	观察频数 f_1	观察频数 f_2	合计	期望频数 e_1	期望频数 e_2
1	f_{11}	f_{12}	$f_{1\cdot}$	$n_1 f_{1\cdot}/N$	$f_{1\cdot} - e_{11}$
2	f_{21}	f_{22}	$f_{2\cdot}$	$n_1 f_{2\cdot}/N$	$f_{2\cdot} - e_{21}$
⋮	⋮	⋮	⋮	⋮	⋮
r	f_{r1}	f_{r2}	$f_{r\cdot}$	$n_1 f_{r\cdot}/N$	$f_{r\cdot} - e_{r1}$
合计	n_1	n_2	N	n_1	n_2

关于自由度的说明。H_0 为真时，近似 χ^2 分布的统计量 Q 分布的自由度是(4.5) 式中独立项的数目。判定 $F_1(x)$ 是否等于 $F_2(x)$，实质是估计 r 组中每一类型的概率，因此，(4.5) 式待估计的参数有 $2r$ 个。由表 4—4 可知，一旦 r 行确定，频数 f_{ij} 项数被确定，期望频数中有一个确定，另一个就被确定。因此(4.5) 式中独立项的数为 $2r - r - 1$，Q 抽样分布的自由度 $df = r - 1$。

二、应用

【例 4.4】　已婚和独身妇女年内没有工作日数的分布是否

不同

已婚女人是否比独身女人更容易请假而从事工作的时间更少。这里没有工作的时间不包括正常的休假，如怀孕，住院治病等。那些离婚、分居、丧偶但身边无子女生活在一起的，均列入独身一组，为验证二者是否不同，随机地分别从两个总体中抽选100人和200人进行调查，结果如表4—5

表4—5　已婚和独身妇女年内无工作日数

无工作日数分组(天)	已婚妇女(人)	独身妇女(人)
0—3	60	130
4—7	21	50
8—11	11	10
12—15	4	6
16—19	2	3
20以上	2	1
合计	100	200

分析：根据所研究的问题，可以建立假设组

H_0：　已婚和独身妇女年内无工作日数分布相同

H_1：　已婚和独身妇女年内无工作日数分布有差异

已婚妇女为第一个样本，独身妇女为第二个样本，x表示没有工作的天数。利用χ^2检验考察是否存在差异。Q统计量计算过程如表4—6。

与单样本的χ^2检验一样，当两个样本中无论哪一个的某一组其期望频数小于5时，应将相邻组合并，同时另一样本相应的两组也要合并，以保证两个样本的组数相等。这种情况下，原来的r组被减少，相应地自由度df也被减少。在这个问题中，“16—19”天和“20及以上”天的期望频数均较小，将其合并，如表4—6的计算那样，天数的分组由6组变为5组，相应地自由度df变为$5-1=4$。在附表Ⅰ中，$\mathrm{df}=4$，$\alpha=0.05$，$\chi^2_{0.05}(4)=9.49$。由于$Q=5.3395$

$< \chi^2_{0.05}(4) = 9.49$，因此，调查数据支持 H_0，表明在显著性水平 $\alpha = 0.05$ 水平，不能拒绝 H_0。已婚和独身妇女在年内没有工作的天数的分布没有什么不同。或者在附表 I 中，查找 H_0 成立时，与 df $= 4, Q = 5.3395$ 相对应的概率 $P > 0.20$，这对显著性水平 $\alpha = 0.05$ 来说，显然已足够大，同样可以得出 H_0 被调查数据支持的结论。

表 4-6　　　　Q 统计量计算表

x	f_1	f_2	$f_{\cdot}$	e_1	e_2	$(f_1-e_1)^2/e_1$	$(f_2-e_2)^2/e_2$
0—3	60	130	190	63.33	126.67	0.1751	0.0875
4—7	21	50	71	23.67	47.33	0.3012	0.1506
8—11	11	10	21	7.00	14.00	2.2857	1.1429
12—15	4	6	10	3.33	6.67	0.1348	0.0673
16—19 20 及以上	2 2 } 4	3 1 } 4	5 3 } 8	2.67	5.33	0.6625	0.3319
合计	100	200	300			3.5593	1.7802
	$Q = 3.5593 + 1.7802 = 5.8395$					df $= 5 - 1 = 4$	

§4.4　两样本的 Kolmogorov-Smirnov 检验

单样本的 K-S 检也可以推广应用于两个独立样本。两样本的 K-S 检验与 χ^2 检验类似，也用于检验总体分布是否相同。

一、基本方法

两个连续总体，具有累积概率分布分别为 $F_1(x)$ 和 $F_2(x)$，要检验两个总体分布是否相同，建立的假设组为

$$H_0:\quad F_1(x) = F_2(x)\qquad \text{对所有 } x$$

$$H_1:\quad F_1(x) \neq F_2(x)\qquad \text{对某个 } x$$

为对假设作出判定，应从两个总体中随机抽选两个独立的样本，数据大小分别记作 m、n。数据的测量层次至少在定距尺度上，若是定序尺度，需能确定两个样本观察值相对差值的大小，两个样

H_1:　两地区青年人的文化程度分布有差异

由于两个地区抽选的样本可视为独立的，且数据是四个以上定序资料组，易于应用 K-S 检验。两个样本的大小不等，因为 A 地区的样本数目较少，作为第一个样本 $m = 236$，则 B 地区为第二个样本，$n = 274$。计算检验统计量 D 的过程如表 4—8。由表可知

表 4—8　　检验统计量 D 的计算表

	绝对频数		累积频数		经验分布函数		
	f_1	f_2	Σf_1	Σf_2	$S_1(x)=\Sigma f_1/m$	$S_2(x)=\Sigma f_2/n$	$\lvert S_1(x)-S_2(x)\rvert$
识字不多或文盲	58	31	58	31	0.2458	0.1131	0.1327
小学	51	46	109	77	0.4619	0.2810	0.1809
初中	47	53	156	130	0.6610	0.4745	0.1865
高中或中专	44	73	200	203	0.8475	0.7409	0.1066
大专	22	51	222	254	0.9408	0.9270	0.0138
大专以上	14	20	236	274	1.0000	1.0000	1.0000
合计	236	274					

$$D = \max|S_1(x) - S_2(x)| = 0.1865$$

由于 m,n 均较大，为大样本，在附表 XI 中需先计算 $\sqrt{N/mn}$ 方可查到合适的 P 值。$m = 236, n = 274, N = m + n = 510$，$\sqrt{N/mn} = 0.0888$。当 $P = 0.010$ 时，临界值是

$$1.63\sqrt{N/mn} = 1.63(0.0888) = 0.1448$$

因为 $D = 0.1865$，大于这个临界值。所以，渐近的近似值是 $P < 0.01$。显然对于显著性水平 $\alpha = 0.05$（或 0.01）P 值够小，因而数据不支持 H_0。检验结果表明，两个地区青年人的文化程度分布存在着明显的差异。

【例 4.6】　城郊县是否比边远县有较低的人口增长速度

表 4—9　　检验统计量的计算表

人口增长速度分组(x)	f_1	f_2	Σf_1	Σf_2	$S_1(x)$	$S_2(x)$	$S_1(x)-S_2(x)$
2.1	1	0	1	0	1/7	0	1/7
2.7	1	0	2	0	2/7	0	2/7
3.1	1	0	3	0	3/7	0	3/7
3.2	1	0	4	0	4/7	0	4/7
3.4	1	0	5	0	5/7	0	5/7
3.5	0	1	5	1	5/7	1/9	38/63
3.7	0	1	5	2	5/7	2/9	31/63
4.8	0	1	5	3	5/7	3/9	24/63
5.3	0	1	5	4	5/7	4/9	17/63
5.6	0	1	5	5	5/7	5/9	10/63
5.8	0	1	5	6	5/7	6/9	3/63
6.8	0	1	5	7	5/7	7/9	− 4/63
7.9	1	0	6	7	6/7	7/9	5/63
8.2	1	0	7	7	1	7/9	2/9
9.3	0	1	7	8	1	8/9	1/9
10.3	0	1	7	9	1	1	0

三、几种检验的比较

两个独立样本的非参数检验，都是检验两个独立样本是否可能来自同一个总体。但是各种检验方法对样本间不同类型差异的敏感程度不一样。若研究的是两个样本是否代表位置(集中趋势)有差异的总体，应选择对这种差异最敏感的检验方法。如Mann-Whitney-Wilcoxon 检验，Kolmogorov-Smirnov 检验(单侧)。在样本容量较大或测量层次较低时，可以采用Mann-Whitney-Wilcoxon 检验，它是专门揭示位置是否有差异的检验。如果样本容量非常小，或者同分秩较多，不便于应用Mann-Whitney-Wilcoxon 检验时，Kolmogorov-Smirnov 检验比 U 检验稍为有效一些。如果研究的是两个样本是否代表任一方面有差异的总体，如位置、离散度、偏斜度等等，可以选用 χ^2 检验、

Kolmogorov-Smirnov 检验(双侧)、Wald-Wolfowitz 检验。当被分析的数据测量层次是定类尺度时，只能采用 χ^2 检验。若被评价的总体是连续分布的，可选用 Wald-Wolfowitz 检验或 Kolmogorov-Smirnov 检验。一般来说，Kolmogorov-Smirnov 检验要比 Wald-Wolfowitz 检验更有效，当数据不满足连续性假定时，它仍然可以适用，只是得到的 P 值将比应得到的稍大些，也就是说犯第 Ⅱ 类错误的概率会稍稍增大。

第五章　k 个相关样本的非参数检验

在参数统计中，检验几个样本是否来自完全相同的总体，采用方差分析或 F 检验。运用 F 检验的假定条件是：样本是从正态分布的总体中独立抽选的；总体具有相同的方差；数据的测量层次至少是定距尺度。当被用来分析的数据不符合这些假定条件，或研究者不希望作这些假设，以便增加结论的普遍性时，不宜采用参数统计的方法，而必须运用非参数方法。

如果 k（等于或大于 3）个样本是按某种或某些条件匹配的，那么 k 个样本称为相关的，否则为独立的。k 个相关和独立样本的差别与两个相关和独立样本之间的差别类似。本章介绍 k 个相关样本的非参数检验。

§5.1　Cochran Q 检验

Cochran Q 检验也译为科库兰检验。它是用以检验匹配的三组或三组以上的频数或比例之间有无显著差异的方法。这种匹配可以用不同形式获得。例如，检验三种不同类型的采访形式对被采访者的有效回答是否有影响，可以抽选一些人，分成 n 组，每组有 3 个匹配的被采访者，要求他们的有关情况相同。每组的 3 名成员被随机地置于 3 种条件之下，即分别接受三种类型的采访，于是，就获得了 3 个匹配的样本，即 $k=3$，每个样本有 n 个观测结果。k 个相关样本也可以采用同一组人，对不同的 k 个条件的反应匹配成样本，这类似于两个相关样本中以研究对象作为自身的对照者。例如，检验几种教学手段对学生掌握知识是否有显著不同，可以随机抽取 n 个学生，让他们先后置于 k 种教学手段之下，再作出评价。这样可以获得 k 个匹配的样本，每个样本有 n 个观测结果。

一、基本方法

若有 k 个相关样本，每个样本有 n 个观测结果，检验 k 个样本间是否有显著差异，可以建立双侧备择，假设组为

H_0： k 个样本间无显著差异

H_1： k 个样本间有显著差异

由于三个及三个以上样本间差异的方向不便于判定，因而，通常只建立双侧备择进行检验。

为对假设作出判定，所分析的数据测量层次为定类尺度即可。获得的数据可排成一个 n 行 k 列的表。如果 H_0 为真，那么将测量结果分为"成功"和"失败"的话，"成功"与"失败"应随机地分布在表中的各行各列。Cochran Q 检验的统计量定义为

$$Q = \frac{(k-1)\left[k\sum_{j=1}^{k}x_j^2 - \left(\sum_{j=1}^{k}x_j\right)^2\right]}{k\sum_{i=1}^{n}y_i - \sum_{i=1}^{n}y_i^2} \qquad (5.1)$$

式中，x_j 是第 j 列的总数，y_i 是第 i 行的总数。由于 Q 统计量的抽样分布近似为自由度 $df = k-1$ 的 χ^2 分布，所以根据自由度 $df = k-1$，给定的显著性水平 α，能够在附表 I 中查找临界值 χ^2_α，若

$Q \geq \chi^2_\alpha$

则在显著性水平 α 下拒绝 H_0，表明样本之间存在着显著差异。相反，则不能拒绝 H_0。

二、应用

【例 5.1】消费者对饮料的爱好是否存在差异

某商店为决定经营饮料的品种、数量，对消费者的爱好进行了一次调查。随机抽取 18 个消费者，请他们对四种饮料：热牛奶、酸奶、果汁、可口可乐的喜好作出评价，凡喜好的记作 1，不喜好记作 0。调查结果如表 5—1。

表 5—1　　消费者对饮料喜好的调查结果

消费者	热牛奶	酸奶	果汁	可口可乐	合计(y_i)
1	1	0	0	1	2
2	0	0	1	0	1
3	0	0	1	1	2
4	1	1	0	0	2
5	1	0	1	0	2
6	0	1	0	0	1
7	0	0	0	1	1
8	0	1	0	0	1
9	0	1	1	0	2
10	1	1	1	0	3
11	0	0	1	0	1
12	0	0	1	0	1
13	1	0	0	1	2
14	1	1	0	0	2
15	1	1	0	0	2
16	0	1	0	0	1
17	1	0	0	1	2
18	0	0	0	1	1
合计(x_j)	8	8	7	6	29

分析：为检验消费者对四种饮料的爱好是否有差异，建立双侧备择，假设组为

H_0：　消费者对四种饮料爱好无差异

H_1：　消费者对四种饮料爱好有差异

由于数据为定类尺度测量，只有“爱好”与“不爱好”两种结果，且是两个以上相关样本，这里是四种饮料，$k=4$，所以选用 Cochran Q 检验。

根据表 5—1 的调查数据，计算 H_0 成立时的统计量 Q。$x_1 = 8$ 表示喜欢第一种饮料热牛奶的总次数，$x_2 = 8$ 是喜欢酸奶的总次数。同样地，$x_3 = 7$，$x_4 = 6$，分别表示消费者喜欢果汁、可口可乐的总次数。$\sum_{j=1}^{4} x_j = 29$ 是所有四种饮料中，消费者表示喜欢的总次数。y_i 是第 i 个消费者喜欢各种饮料的次数。$\sum_{i=1}^{18} y_i = 29$，是各个消费者对四种饮料表示喜欢的总次数。$\sum_{j=1}^{k} x_j$ 表示按样本数计算的消费者喜欢的总次数，而 $\sum_{i=1}^{n} y_i$ 表示按观察对象即消费者或说按样品数计算的对各种饮料喜欢的总次数。这两个总和应相等，即有 $\sum_{j=1}^{k} x_j = \sum_{i=1}^{n} y_i$。统计量 Q 正是用于说明按样本数计算的总次数与按样品数计算的总次数的符合程度。按(5.1)式，

$$
\begin{aligned}
Q &= \frac{(4-1)[4(64+64+49+36)-29^2]}{4(29)-[9(2^2)+8(1^2)+3^2]} \\
&= \frac{3[852-841]}{116-53} \\
&= 0.5238
\end{aligned}
$$

根据给定的显著性水平 $\alpha = 0.05$，自由度 $\mathrm{df} = 4-1 = 3$，查附表 I，得到临界值 $\chi_\alpha^2 = 7.82$。显然，$Q = 0.5238 < \chi_\alpha^2 = 7.82$。因而，调查数据在 5% 的显著性水平上不能拒绝 H_0，即消费者对四种饮料的爱好没有显著差异。

【例 5.2】 三种不同教学方法的效果是否有显著差异

三种不同教学方法：电视教学、课堂讲授、课堂讨论，对学生掌握知识的效果是否有所不同。为检验这一问题，抽选部分学生分为 18 组，每组 3 名匹配的学生，他们的有关情况类似。各组中 3 名学生被随机地置于 3 种条件下，即随机地指定接受某种教学方法。实施不同教学方法后进行测验，成绩合格为有效，记作 1；成绩不合格为无效，记作 0。结果如表 5—2。

分析：学生的考试成绩是定距尺度测量，这里将其转化为合

格、不合格两类，则视为定类尺度。合格即教学方法有效为 1，不合格 为教学方法无效，记作 0。接受三种不同教学方法的学生在每一组是匹配的，即构成 3 个相关样本，$k=3$。检验三种教学方法的效果是否存在差异，建立的假设组为

H_0： 三种教学方法的效果无显著差异

H_1： 三种教学方法的效果有显著差异

表 5—2　　实施不同教学方法的学生成绩

学生组	电视教学	课堂讲授	课堂讨论	合计(y_i)
1	0	0	0	0
2	0	1	1	2
3	0	1	0	1
4	0	0	0	0
5	1	0	1	2
6	0	1	1	2
7	0	1	1	2
8	0	1	0	1
9	1	0	1	2
10	0	0	0	0
11	0	1	1	2
12	0	1	1	2
13	0	1	1	2
14	0	1	1	2
15	0	1	1	2
16	1	1	1	3
17	0	0	1	1
18	0	1	1	2
合计(x_j)	3	12	13	28

由于是定类尺度测量的数据，相关样本数目大于 2，因此，宜

采用 Cochran Q 检验。

利用表 5—2 的数据计算检验统计量 Q

$$Q = \frac{(3-1)[3(3^2+12^2+13^2)-28^2]}{3(28)-[11(2^2)+3(1^2)+3^2]}$$

$$= \frac{2[3(322)-784]}{84-56}$$

$$= 13$$

给定显著性水平 $\alpha = 0.05$，$df = 3-1 = 2$，查附表 I 中相应临界值 $\chi^2_\alpha = 5.99$。显然，$Q = 13 > \chi^2_\alpha = 5.99$，在 5% 的显著性水平上调查数据拒绝 H_0，表明三种不同教学方法的效果有显著差异。最后的判定，还可以采用这种方法。查附表 I，对于自由度 $df = 3-1 = 2$，在 H_0 为真时，$Q \geqslant 13$ 出现的概率 P 接近于 0.01。这一概率显然小于显著性水平 $\alpha = 0.05$，也小于显著性水平 $\alpha = 0.01$，因此 Q 值位于否定域内，调查结果拒绝 H_0。

运用 Cochran Q 检验时应注意，只有当行数 n 不太小时，Q 的抽样分布才近似于 $df = k-1$ 的 χ^2 分布。但是，n 的最小数值目前并没有明确的说明，使用者采用时视具体问题而定。Cochran Q 检验适用于定类尺度测量的数据，其它测量层次的数据也可以运用，但要象例 5.2 那样，转化为两类，但这样做可能浪费数据中包含的信息。因此，Cochran Q 检验一般只用于定类尺度的数据。

§5.2 Friedman 检验

Friedman 检验亦称佛利得曼的 χ^2 检验。或佛利得曼双向评秩方差分析。它是对 k 个样本是否来自同一总体的检验。k 个样本是匹配的，实现匹配的方法与前面类似。可以是 k 个条件下同一组受试者构成，即受试对象作为自身的对照者，也可以将受试者分为 n 个组，每组均有 k 个匹配的受试者，随机地将 k 个受试者置于 k 个条件之下形成。在不同受试者匹配的样本中，应尽量使不同受试者的有关因素匹配即相似。

一、基本方法

与 Cochran Q 检验相似，Friedman 检验也是用来检验各个样本所得的结果在整体上是否存在显著差异。因此建立的也是双侧备择，假设组为

H_0: k 个样本间无显著差异

H_1: k 个样本间有显著差异

为对假设作出判定，所分析的数据应是定序尺度测量。获得的数据排成一个 n 行 k 列的表，行代表不同的受试者或匹配的受试小组，列代表各种条件。由于是定序尺度测量的数据，因此，可以对每一行的观测结果分别评秩，即评等级，等级 1 是最小的，依次排序，秩从 1 到 k。如果 H_0 为真，那么每一列中秩的分布应该是随机的，即各个秩出现在所有列中的频数应几乎相等，也就是说各列的秩和应该大致相等。Friedman 检验定义的统计量为 χ_r^2。

$$\chi_r^2 = \frac{12}{nk(k+1)}\sum_{j=1}^{k}R_j^2 - 3n(k-1) \qquad (5.2)$$

式中，R_j 是第 j 列的秩和，即等级和。χ_r^2 的抽样分布在 n、k 不太小时，近似于自由度 $\mathrm{df}=k-1$ 的 χ^2 分布。因此，在附表 I 中，可以根据给定的显著性水平 α，自由度 $\mathrm{df}=k-1$ 查得 H_0 为真时，相应的临界值 χ_α^2。若 $\chi_r^2 \geqslant \chi_\alpha^2$，则在 α 水平上拒绝 H_0，否则不能拒绝 H_0。

二、应用

【例 5.3】 三种不同教学方法的效果是否有显著差异

三种不同教学方法同例 5.2，抽选的学生也分为 18 组，每组 3 名匹配的学生，其有关情况类似。各组中 3 名学生被随机地安排接受某种教学方法。实施不同教学方法后，进行测验，按成绩高低对 3 名匹配学生的成绩排列等级即评秩，结果如表 5—3。

分析：这个问题与例 5.2 类似，也是检验三种教学方法的效果有无差异，因而应建立双侧备择，假设组为

H_0: 三种学方法的效果无显著差异

H_1: 三种教学方法的效果有显著差异

表 5—3　实施不同教学方法的学生成绩等级

学生组	电视教学	课堂讲授	课堂讨论
1	1	3	2
2	1	2	3
3	2	3	1
4	3	2	1
5	2	1	3
6	1	3	2
7	1	2	3
8	2	3	1
9	2	1	3
10	2	1	3
11	1	3	2
12	1	3	2
13	1	2	3
14	1	3	2
15	1	2.5	2.5
16	1	2	3
17	1	2	3
18	1	2	3
合计(R_j)	25	40.5	42.5

由于数据的测量已转化为定序尺度，且是两个以上相关样本，故可以采用 Friedman 检验。

根据表 5—3 的数据，按(5.2)式计算检验统计量 χ_r^2。

$$\chi_r^2 = \frac{12}{nk(k+1)}\sum_{j=1}^{k} R_j^2 - 3n(k+1)$$

$$= \frac{12}{18(3)(3+1)}(25^2 + 40.5^2 + 42.5^2) - 3(18)(3+1)$$
$$= 226.8 - 216 = 10.8$$

给定显著性水平 $\alpha = 0.05$，自由度 $df = k - 1 = 2$，查附表 I 中 H_0 成立时相应的临界值 $\chi^2_\alpha = 5.99$。显然，$\chi^2_r = 10.8 > \chi^2_\alpha = 5.99$，因此数据在 5% 的显著性水平上拒绝 H_0，三种教学方法的效果有显著差异。

表 5—3 中，第 15 组接受课堂讲授和课堂讨论方法的学生测验成绩相同，因此排序时，取秩 2 和 3 的平均值，均记为 2.5。以平均秩替代同分，不影响这一检验的有效性。

【例 5.4】 四部分技术训练的有效性有无差异

表 5—4 学员受训后检测的成绩

学员编号	技术训练 I	技术训练 II	技术训练 III	技术训练 IV
1	10	3	6	8
2	2	5	9	4
3	4	10	3	8
4	6	3	10	4
5	3	4	10	6
6	5	4	6	7
7	7	10	6	5
8	6	10	3	5
9	10	5	7	6
10	8	9	7	6
11	5	4	2	6
12	3	5	4	7
13	4	5	10	9
14	6	5	8	10

某田径队对新入队的学员要进行四个部分的技术训练，以提高学员的身体素质。为检验这四个部分的技术训练计划是否确实

$\doteq 210.7714 - 210 = 0.7714$

在附表 I 中，查找与显著性水平 $\alpha = 0.05$，自由度 $df = k - 1 = 3$ 相对应的临界值 $\chi_\alpha^2 = 7.82$。显然 $Q = 0.7714 < \chi_\alpha^2 = 7.82$，调查结果在 5% 的显著性水平上不能拒绝 H_0，表明四个技术训练的有效性没有显著差异。

三、Cochran Q 检验与 Friedman 检验

这两个检验都用于 k 个相关样本是否可能来自同一个总体的检验。但对数据测量层次的要求不同。当数据为定类尺度测量，只能运用 Cochran Q 检验。因为，这一检验对于定类尺度或仅分为两类的定序尺度测量数据是极为有效的。若数据测量层次至少为定序尺度时，应优先选用 Friedman χ_r^2 检验。因为若将定序尺度转换为定类尺度，而采用 Cochran Q 检验，可能会浪费数据包含的信息。

第六章　k 个独立样本的非参数检验

在统计的分析研究中，常常需要确定 k 个独立样本是否来自同一总体或从 k 个相同总体中抽取。通常用以检验这一问题的参数方法是单向方差分析和 F 检验。运用参数方法的假设如第五章所述，而实际分析的数据往往并不具备这些条件或不必要加以限制而使结论更有普遍意义。这时，不能运用参数方法，而只能采用非参数方法。本章介绍几种 k 个独立样本的非参数检验方法。

§6.1　Kruskal-Wallis 检验

Kruskal-Wallis 检验亦有译为克拉夏尔－瓦里斯检验，或简称为克氏检验。它是两个独立样本 Mann-Whitney-Wilcoxon 检验的一种推广。

一、基本方法

若有 k 个总体，各自的连续累积分布函数为 $F_1(x),F_2(x),\cdots,F_k(x)$，那么 Kruskal-Wallis 检验的一般零假设为

$$H_0:\quad F_1(x)=F_2(x)\cdots=F_k(x)\qquad \text{对所有 } x$$

如果在研究总体是否相同时，偏重于考察位置参数，并且位置参数采用各个总体的中位数，即么，H_0 等价于 k 个总体的中位数相等。若仍以 $M_1,M_2,\cdots,M_k$ 代表 k 个总体的中位数，则 Kruskal-Wallis 检验建立的假设组为

$$H_0:\quad M_1=M_2=\cdots=M_k$$

$$H_1:\quad M_j(j=1,2,\cdots,k)\text{ 中至少有两个不相等}$$

这里的备择对于 $k>2$ 时不存在单侧备择的配对，因为对于 $M_j(j=1,2\cdots,k)$ 来说，有 $k!=k(k-1)\cdots(1)$ 种不同的有序排列，这不便于进行检验。

为对假设作出判定，需要的数据是 k 个独立的随机样本，其大

种方法有效性是否有差异，可以建立假设组为

H_0：　$M_1 = M_2 = M_3 = M_4$

H_1：　$M_j(j = 1,2,3,4)$ 中至少有两个不等

由于数据是定序尺度测量，有两个以上独立样本，因此可以采用 Kruskal-Wallis 检验。根据表 6—1 的数据，按(6.2)式计算检验统计量 H。

$$H = \frac{12}{40(40+1)}\left(\frac{260^2}{10} + \frac{122^2}{10} + \frac{90^2}{10} + \frac{384^2}{10}\right) - 3(40+1)$$

$$\doteq 31.89$$

在附表 1 中，与 $df = k - 1 = 3$，显著性水平 $\alpha = 0.05$ 相对应的临界值 $\chi^2_\alpha = 7.82$。显然 $H = 31.89 > \chi^2_\alpha = 7.82$。数据在 5% 的显著性水平上拒绝 H_0，表明四种不同治疗方法对精神错乱的有效性存在显著差异。

【例 6.2】　四种培训方案的有效性是否存在显著差异

为培训大学生志愿者为社区服务，设计了 4 种培训方案，记作 A、B、C、D。将报名的 30 名大学生随机地分为 4 组，分别接受不同培训。训练一周后，按规定的要求考试，评定的成绩如表 6—2

表 6—2　　志愿者培训后考试成绩

培训方案 A	培训方案 B	培训方案 C	培训方案 D
60	72	61	63
75	52	85	58
62	68	78	65
76	82	66	71
73	74	70	84
98	64	59	77
86	87	69	80
		79	89

分析：对志愿者实施四种培训方案后测试的成绩是定距尺度

的分数，可以将其按从小到大的次序给予适当的等级。将所有 30 名大学生志愿者的成绩混合排列，逐个评秩，结果如表 6—3。由于数据可以评秩，样本为 4 个独立样本，可以采用 Kruskal-Wallis 检验。以各样本的中位数是否相等作为判定四种方案有效性是否存在差异的标志。建立的假设组为

表 6—3　　志愿者成绩的等级

	培训方案 A	培训方案 B	培训方案 C	培训方案 D
	4	15	5	7
	18	1	26	2
	6	11	21	9
	19	24	10	14
	16	17	13	25
	30	8	3	20
	27	28	12	23
			22	29
合计(R_j)	120	104	112	129

H_0:　$M_1 = M_2 = M_3 = M_4$

H_1:　$M_j(j = 1,2,3,4)$ 中至少有两个不等

利用表 6—3 的秩次和，按(6.2)式计算统计量 H，得到

$$H = \frac{12}{30(30+1)}\left(\frac{120^2}{7} + \frac{104^2}{7} + \frac{112^2}{8} + \frac{129^2}{8}\right) - 3(30+1)$$

$$\doteq 93.5537 - 93$$

$$= 0.5537$$

附表 I 中，显著性水平 $\alpha = 0.05$，$\mathrm{df} = k - 1 = 3$ 时，临界值 $\chi^2_\alpha = 7.82$。显然 $H = 0.5537 < \chi^2_\alpha = 7.82$，数据在 5% 的显著性水平上不能拒绝 H_0，表明四种培训方案的有效性没有什么显著差异。最后的判定也可以通过 P 值与显著性水平 α 比较得到。附表 I 中，与 $\mathrm{df} = k - 1 = 3$、$H = 0.5537$ 相对应的概率是 $P > 0.90$。显然 H_0 为真时，$H = 0.5537$ 出现的概率远远大于显著性水平 $\alpha = 0.05$。

因此，同样可以作出结论，调查数据在 5% 的显著性水平上不能拒绝 H_0。

三、同分的处理

例 6.2 中的每个志愿者成绩均不相同，因而没有同分问题，但在实际中，往往会出现评分相同的情况。如果在两个或两个以上的评分之间出现同分时，每一个评分的秩都记作这些同分秩的平均值。由于出现同分会对统计量 H 有影响，因而计算 H 值时，应进行校正。校正系数为

$$1-\frac{\Sigma u^3-\Sigma u}{N(N^2+1)} \tag{6.3}$$

式中，u 是相同评分的观察值数目，如学员考试成绩有 2 个 62 分，则 $u=2$；还有 4 个 78 分，则 $u=4$ 等等。计算 H 值时，利用 (6.2) 式除以 (6.3) 式，得到的是校正的 H 值。经过校正以后 H 值比校正前要大。如果未校正时，计算结果就能拒绝 H_0，那么校正后将在更加苛刻的显著性水平上拒绝 H_0，因为与较大的 H 值相对应的概率 P 值将更小。在大多数情况下，这一校正常可忽略。根据 Kruskal 和 Wallis 在 1952 年的著作中证明，当同分的观察值数目占观察值总数目的比例不到 25% 时，校正后的概率仅仅改变百分之十几。一般情况，校正因子的大小取决于 u 值的大小，即同分的数目和同分观察值数目占观察值总数的百分比。

【例 6.3】 三种不同教学方法的有效性是否有显著差异

某大学制定三种不同的教学方法：大班讲授、小组讲授、小组讨论。为检验三种方法对学生掌握知识的有效性是否相同，进行了一次试验。选取二年级大学生 50 名，随机地分为三组，分别接受三种不同方法教学。由同一教师按不同方法分别讲授同一方面的知识，规定的内容讲授完后，对学生进行统一考试，成绩如表 6—4。

表 6-4　　学生考试成绩

大班讲授组(Ⅰ)	小组讲授组(Ⅱ)	小组讨论组(Ⅲ)
62　73　56	73	84
56　79　48	78	86
62　89　64	92	98
84　98　72	86	72
90　92　78	84	69
48　52　84	69	79
49　54　86	73	86
64　84　92	92	84
69　82　98	98	70
72　69　62	81	90

分析：学生成绩为定距尺度测量，但为了避免作出某些假设，以使结论更具普遍性，所以不准备采用参数检验方法，而选用非参数检验。由于三种不同教学方法是独立的，故应采用 k 个独立样本的统计检验。对于三组学生成绩集中趋势的一个很好的度量指标是中位数，成绩可以由小到大排序给出等级，因此能够采用 Kruskal-Wallis 检验。建立假设组为

H_0:　$M_1 = M_2 = M_3$

H_1:　$M_j (j = 1,2,3)$ 中至少有两个不等

若用文字描述为

H_0:　接受不同教学方法的学生平均成绩没有显著差异

H_1:　接受不同教学方法的学生平均成绩不完全相同

为采用 Kruskal-Wallis 检验对假设作出判定，将表 6—4 中的所有学生成绩排序，最低分秩评为 1，最高分秩评为 50。由于 50 名学生中有不少是同分，采用相应秩的简单算术平均数作为同分的平均秩，得到表 6—5。用表中数据，按(6.2)式计算得到

$$H = \frac{12}{50(50+1)}\left(\frac{649.5^2}{30} + \frac{318.5^2}{10} + \frac{307^2}{10}\right) - 3(50+1)$$

$\doteq 158.2626 - 153$

$= 5.2626$

表 6—5　　　　　　**学生考试成绩的等级**

	Ⅰ			Ⅱ	Ⅲ
	9	22	6.5	22	32.5
	6.5	26.5	1.5	24.5	37.5
	9	40	11.5	44.5	48.5
	32.5	48.5	19	37.5	19
	41.5	44.5	24.5	32.5	14.5
	1.5	4	32.5	14.5	26.5
	3	5	37.5	22	37.5
	11.5	32.5	44.5	44.5	32.5
	14.5	29	48.5	48.5	17
	19	14.5	9	28	41.5
合计(R_i)			649.5	318.5	307

附表 Ⅰ 中，$df = k - 1 = 2$，$H = 5.2626$ 出现的概率 P 在 0.05 与 0.10 之间。若显著性水平 $\alpha = 0.05$，则数据不能拒绝 H_0；而显著性水平 $\alpha = 0.10$，则数据拒绝 H_0。由于学生成绩中同分较多，因而应采用校正的 H。计算同分的观察值数目，即 u 和 u^3，计算过程列于表 6—6 中。由于 $\Sigma u^3 - \Sigma u = 601 - 43 = 558$，所以校正因子为

$$1 - \frac{\Sigma u^3 - \Sigma u}{N(N^2 - 1)} = 1 - \frac{558}{50(50^2 - 1)} \doteq 0.9955$$

校正后的统计量 H 为

$$H = \frac{5.2626}{0.9955} \doteq 5.2864$$

这一结果与校正前的 H 值相差不多。对于显著性水平 $\alpha = 0.05$，$df = k - 1 = 2$，H_0 为真时的临界值为 $\chi_\alpha^2 = 5.99$。$H = 5.2864 < \chi_\alpha^2 = 5.99$，数据在 5% 的显著性水平上不能拒绝 H_0，表明接受不同教学方法的学生平均成绩没有显著差异。

表 6－6　　　　同分的观察值数目计算

同分的观察值	u	u^3
48	2	3
56	2	8
62	3	27
64	2	8
69	4	64
72	3	27
73	3	27
78	2	8
79	2	8
84	6	216
86	4	64
90	2	8
92	4	64
98	4	64
合计	43	601

§6.2　k 个样本的 χ^2 检验

两个独立样本的 χ^2 检验可以直接推广到 k 个独立样本，用来检验 k 个样本之间差异的显著性。

一、基本方法

k 个独立样本 χ^2 检验与两个独立样本的基本方法类似。零假设是 k 个样本来自同一总体，或来自一些相同的总体。样本可以是 k 个频数或 k 个比例。每一样本都可以分成 r 组，因此数据可以排成一个 $k\times r$ 的表。若以 f_{ij} 表示第 i 行第 j 列的实际频数或比例，e_{ij} 表示与其相应的理论频数或比例，那么，检验统计量 Q 为

$$Q = \sum_{i=1}^{r} \sum_{j=1}^{k} \frac{(f_{ij} - e_{ij})^2}{e_{ij}} \qquad (6.4)$$

以 $f_{i.}$ 表示第 i 行的频数和或比例和，k 个样本观察值的数目分别为 $n_1 = \Sigma f_1, n_2 = \Sigma f_2, \cdots, n_k = \Sigma f_k$，总数目 $N = n_1 + n_2 + \cdots + n_k$，任一样本观察值数目记作 $n_j (j = 1, 2, \cdots, k)$，那么，理论频数 e_{ij} 可以由下式计算

$$e_{ij} = n_j (f_{i.}/N) \qquad (6.5)$$

当 H_0 为真时，Q 统计量的抽样分布近似于自由度 $\mathrm{df} = (k-1)(r-1)$ 的 χ^2 分布。根据给定的显著性水平 α，自由度 $\mathrm{df} = (k-1)(r-1)$，在附表 I 中，可以查到 H_0 成立时的临界值 χ^2_α。若 $Q \geqslant \chi^2_\alpha$，则数据在 α 水平上拒绝 H_0；$Q < \chi^2_\alpha$，数据在 α 水平上不能拒绝 H_0。

二、应用

【例 6.4】 收听体育广播兴趣不同的人，参加体育活动的情况是否也不相同

1990 年秋的《亚运会》调查，在全国范围抽选 12 个省的 2162 人(原抽选 2211 人，其中 49 人未回答)，按收听体育广播的兴趣程度分为：很不喜欢、不喜欢、无所谓、喜欢、非常喜欢五类，各类人员参加体育活动情况如表 6—7。

表 6—7 各类人员参加体育活动情况的人数

	很不喜欢	不喜欢	无所谓	喜欢	非常喜欢	合计
不参加	15	96	641	290	57	1102
偶尔参加	3	15	218	204	49	489
经常参加	4	15	105	113	32	269
天天参加	3	14	119	108	58	302
合计	25	140	1086	715	196	2162

资料来源：柯惠新等：《调查研究中的统计分析》P262，北京广播学院出版社，1992。

分析：按收听体育广播兴趣划分的五种类型是相互独立的，抽选的 2162 人随机地分为这五类，因此，$k = 5$ 是独立样本，应采用 k

$= 141.7795$

若给定显著性水平 $\alpha = 0.05$，由 $df = (k-1)(r-1) = (5-1)(4-1) = 12$，查附表 I，临界值 $\chi^2_\alpha = 21.03$。因为 $Q = 141.7795 > \chi^2_\alpha = 21.03$，所以数据在5%的水平上拒绝 H_0，表明收听体育广播兴趣不同的人，参加体育活动的情况也不同，即收听体育广播的兴趣对参加体育活动有影响。

三、k 个比例相等性检验

若所研究的 k 个独立样本是 k 个比例，那么，对 k 个独立样本是否来自同一总体，或是否来自 k 个相同总体的检验，实际是对 k 个比例相等性的检验。k 个比例分别记作 $P_1, P_2, \cdots, P_k$，则建立的假设组为

$$H_0: \quad P_1 = P_2 = \cdots = P_k$$

$$H_1: \quad P_j (j = 1,2,\cdots,k) \text{ 中至少有两个不等}$$

为了对假设作出判定，所需要的数据是定类尺度测量的。k 个样本的数据个数分别为 $n_1, n_2, \cdots, n_k$ 那么第 j 个样本观察值数目记作 $n_j (j = 1,2,\cdots,k)$。k 个样本的实际频数分别记作 $f_1, f_2, \cdots, f_k$，则 k 个样本的平均比例 $\overline{P}$ 为

$$\overline{P} = \Sigma f_j / \Sigma n_j \qquad (6.6)$$

由(6.6)式可以得到第 j 个样本的期望频数为 $n_j\overline{P}$。若 H_0 为真，那么实际频数 f_j 与相应的期望频数 $n_j\overline{P}$ 应该相等。所以实际频数 f_j 与期望频数 $n_j\overline{P}$ 的偏差可以作为度量 k 个比例是否相等的一个指标。为检验 k 个比例是否相等定义的统计量为 Q。其计算公式为

$$Q = \sum_{j=1}^{k} \frac{(f_j - n_j\overline{P})^2}{n_j\overline{P}(1-\overline{P})} \qquad (6.7)$$

(6.7)式也可以写成

$$Q = \frac{1}{\overline{P}(1-\overline{P})} \sum_{j=1}^{k} n_j \left(\frac{f_j}{n_j} - \overline{P}\right)^2 \qquad (6.8)$$

统计量 Q 的抽样分布近似于自由度 $df = k-1$ 的 χ^2 分布。附表 I 中 χ^2 分布的右尾概率是相应的 P 值。对假设判定的准则和 χ^2 检验

的一般准则相同。

【例 6.4】 三个区域房屋销售的比例是否有差异

某房地产公司准备在一地区购买房屋并出售，由于资金有限，不能在这一地区大量购买。决策者准备先从期望销售比例最大的区域开始，获得利润后再开辟别的区域。为检验这个地区的三个区域房屋销售比例是否有显著不同，随机地从每个区域抽选容量为 100 的样本调查。抽样框是已发出的付款通知书，即已被人认购的房屋。调查结果如下

区　　域	Ⅰ	Ⅱ	Ⅲ
已销售数目	40	25	20
发通知数目	100	100	100

分析：决策者要在三个区域中选择首先开始销售的区域，就是要确定三个区域的房屋销售比例是否相同，也就是检验三个区域已销售房屋数目占发通知书数目的比重是否有显著差异。若有差异，哪个区域的比例最高，应从哪个区域开始。若三个区域的房屋销售比例分别记作 P_1, P_2, P_3，建立的假设组为

H_0: $P_1 = P_2 = P_3$

H_1: P_1, P_2, P_3 不完全相等

假定这三个区域相对于居住单位是均匀的，3 个独立样本的比例是否相等的检验，可以采用 χ^2 检验。由于调查数据为

$f_1 = 40 \qquad f_2 = 25 \qquad f_3 = 20$

$n_1 = 100 \qquad n_2 = 100 \qquad n_3 = 100$

根据(6.6) 式有

$$\overline{P} = \Sigma f / \Sigma n = (40 + 25 + 20)/300 \doteq 0.2833$$

于是三个样本的期望频数分别为

Ⅰ： $n_1\overline{P} = 28.33$

Ⅱ： $n_2\overline{P} = 28.33$

Ⅲ: $n_3\overline{P} = 28.33$

由(6.7)式可以计算得到检验统计量为

$$Q = \frac{(40-28.33)^2}{(28.33)(0.7167)} + \frac{(25-28.33)^2}{(28.33)(0.7167)}$$
$$+ \frac{(20-28.33)^2}{(28.33)(0.7167)}$$
$$\doteq 6.7045 + 0.5461 + 3.4175$$
$$= 10.6681$$

根据给定的显著性水平 $\alpha = 0.05$，自由度 $\mathrm{df} = k - 1 = 2$，在附表Ⅰ中查找相应的临界值 $\chi^2_\alpha = 5.99$。因为 $Q = 10.6681 > \chi^2_\alpha = 5.99$，表明数据在5%的显著性水平上拒绝 H_0，三个区域的销售比例显著不同。若以自由度 $\mathrm{df} = k - 1 = 2$，在附表Ⅰ中查找 $Q = 10.6681$ 时的近似概率，P 在0.001和0.01之间。近似的 $P < 0.01$，表明数据不支持 H_0，三个区域房屋的销售比例显著不同。由于第Ⅰ个区域的实际销售比例最高 $P_1 = f_1/n_1 = 0.40 > P_2 = 0.25 > P_3 = 0.20$ 故应从第一个区域开始实施购房销售计划。

第七章　两个样本的相关分析

所谓相关，是指两组或两组以上观察结果之间的连带性或联系。换句话说，也就是各组观察结果所反映的特性之间有关系。如几个亲生兄弟间的智商与出生顺序有关系，受教育程度与性别有关系等等。在实际问题的研究中，人们常常想知道两组或两组以上的观察结果是否有联系，同时也想知道联系的程度如何。前面的统计检验能够在一定的显著性水平上，确定各组观察值的关系是否存在，本章开始及以后两章，将介绍如何测定各组观察值关系的强度。本章介绍两个样本相关的测定方法。除了相关程度测定方法外，还将介绍判定总体是否存在真实相关关系的相关系数显著性检验。

§7.1　等级相关

等级相关(Rank Correlation)也称作级序相关，用于两个至少是定序尺度测量的样本间相关程度的测定。

一、基本方法

两个样本 X,Y，其观察数据可以配对为(x_1,y_1)，(x_2,y_2)，…，(x_n,y_n)。将 $x_1,x_2,\cdots,x_n$ 排序后评秩，其秩记作 U，与 x_i 相对应的秩为$U_i(i=1,2,\cdots,n)$；同样，$y_1,y_2,\cdots,y_n$ 排序后评秩，秩记作 V，与 y_i 相对应的秩为 $V_i(i=1,2,\cdots,n)$。这样得到的 n 对秩(U_1,V_1)，(U_2,V_2)，…，(U_n,V_n) 可能每一对完全相等，也可能不等。由于每一样本都是n 个数据评秩，因此U_i与V_i的取值都是从 1 到 n。X、Y 的秩可能完全一致，即对于所有的 i 来说，有 $U_i=V_i$，表 7—1 是完全一致的评秩结果。X、Y 的秩可能完全相反，表 7—2 是完全相反的评秩结果。如果 X、Y 完全相关，应该对于所有的 i 有 $U_i=V_i$，即 $U_i-V_i=0$。因此，U_i 与 V_i 之差可以用来度量 X、Y 的相关

程度。定义

表 7—1　　完全一致的评秩

X 的秩	Y 的秩
1	1
2	2
⋮	⋮
$n-1$	$n-1$
n	n

表 7—2　　完全相反的评秩

X 的秩	Y 的秩
1	n
2	$n-1$
⋮	⋮
$n-1$	2
n	1

$$D_i = U_i - V_i$$

则两组秩完全相关时，$D_i(i=1,2,\cdots,n)$ 应该为零。D_i 越大，X、Y 之间的相关越不完全。但由于 D_i 可正可负，直接用 $\sum\limits_{i=1}^{n} D_i$ 测度相关，会出现正负 D_i 抵消，而不能真实反映 U_i 与 V_i 差值的大小，所以宜采用 $\sum\limits_{i=1}^{n} D_i^2$，即

$$\sum_{i=1}^{n} D_i^2 = \sum_{i=1}^{n} (U_i - V_i)^2 \tag{7.1}$$

(7.1)式的这个秩差值平方和的大小既受到 n 的多少的影响，又受到两组秩不一致程度的影响，因此，采用相对的测量指标有利于说明 X、Y 的相关程度。因为 $\sum\limits_{i=1}^{n} D_i^2$ 的最大值反映 X、Y 完全不相关的情况，所以，用(7.1)式除以 ΣD_i^2 的最大值，可用来评价 X、Y 之间秩的差值是否与完全不相关时接近。若实际计算的 ΣD_i^2 与 X、Y

完全不相关情况下的$\Sigma(U_i - V_i)^2$接近，那么两个样本的相关程度较低；若实际计算的ΣD_i^2与ΣD_i^2最大值的比越小，则两个样本的相关程度越高。ΣD_i^2的最大值即X、Y间完全不相关情况下的秩差值平方和，可以根据表7—2所列的数据计算。因为这是X、Y完全不相关的评秩结果。ΣD_i^2的最大值为

$$
\begin{aligned}
&(n-1)^2 + [(n-1)-2]^2 + \cdots + [2-(n-1)]^2 \\
&+ (1-n)^2 \\
&= 2[(n-1)^2 + (n-3)^2 + \cdots] \\
&= n(n^2-1)/3 \qquad (7.2)
\end{aligned}
$$

(7.2)式的中括号内最后一项，当n为奇数时是$[(n+1)/2+1] - [(n+1)/2-1]$的平方，即为$2^2$；$n$为偶数时是$[n/2-(n/2-1)]^2 = 1^2$。

(7.1)式除以(7.2)式得到

$$\frac{\Sigma D_i^2}{n(n^2-1)/3} = \frac{3\Sigma D_i^2}{n(n^2-1)} \qquad (7.3)$$

(7.3)式的取值从0到1。表7—1数据计算的(7.3)式值为0，表7—2数据计算的(7.3)式值为1，即X、Y的秩完全一致时，(7.3)式的值为0，X、Y的秩完全不一致时，(7.3)式的值为1。

测度两个样本等级相关程度可以象参数方法一样，定义等级相关系数作为标准。斯皮尔曼的等级相关系数(Spearman coefficient of rank correlation)是测定两个样本相关强度的重要指标。其计算公式为

$$R = 1 - \frac{6\Sigma D_i^2}{n(n^2-1)} \qquad (7.4)$$

斯皮尔曼相关系数也写为r_s，在右下标注以s是为表明这个相关系数r不是积矩相关的简单相关系数，而是等级相关的Spearman相关系数。由于(7.4)式与(7.3)式不同，所以，R的取值从-1到$+1$，$|R| = 1$表明X、Y完全相关，$R = +1$为完全正相关，$R = -1$为完全负相关。$|R|$越接近于1，表明相关程度越高，反之，$|R|$越接近于零，表明相关程度越低，$R = 0$为完全不相关。$R > 0$为正相

关，$R<0$ 为负相关。通常认为 $|R|>0.8$ 为相关程度较高。

二、应用

【例 7.1】 两个裁判员对参赛歌手评分的相关分析

表 7—3 是某次歌手大赛两名裁判员对 10 名参赛歌手的评分等级，试分析两位裁判员评分的相关程度。

表 7—3　　两个裁判员的评分等级

参赛歌手编号	U	V	$D=U-V$	D^2
1	1	5	−4	16
2	2	3	−1	1
3	5	9	−4	16
4	9	6	3	9
5	4	8	−4	16
6	6	4	2	4
7	3	2	1	1
8	7	1	6	36
9	10	7	3	9
10	8	10	−2	4
合计				112

分析：由于表 7—3 所给数据为评分等级，两个定序数据间的相关程度测定可以采用斯皮尔曼秩相关系数。根据(7.4)式计算得到

$$
\begin{aligned}
R &= 1-\frac{6\Sigma D_i^2}{n(n^2-1)} \\
&= 1-\frac{6(112)}{10(10^2-1)} \\
&= 1-0.6788 \\
&= 0.3212
\end{aligned}
$$

$R=0.3212<0.8$，所以两个裁判员对 10 名参赛歌手评分的

表 7-7　经济水平与卫生水平的秩次

区县编号	X 的秩次(U)	Y 的秩次(V)	$D=U-V$	D^2
1	1	14	-13	169
2	2.5	3.5	-1	1
3	2.5	2	0.5	0.25
4	4	1	3	9
5	5	5	0	0
6	6	6	0	0
7	7	11	-4	16
8	8	3.5	4.5	20.25
9	9.5	9	0.5	0.25
10	9.5	10	-0.5	0.25
11	11	16.5	-5.5	30.25
12	12.5	8	4.5	20.25
13	12.5	15	-2.5	6.25
14	14	7	7	49
15	15	12	3	9
16	16	13	3	9
17	17	18	-1	1
18	18	16.5	1.5	2.25
19	19	20	-1	1
20	20	19	1	1
21	21	21	0	0
22	22	22	0	0
23	23	24	-1	1
24	24	23	1	1
合计				347.00

四、R 的显著性检验

利用(7.4)式或(7.6)式计算的 R 值，是抽自两个总体的样本数据计算的结果，从这一相关系数的大小，可猜测总体的秩相关系数是否与零有显著差异，但是否为真，应进行假设检验。对 R 的显著性检验正是为了回答这一问题。检验可以仅研究两个总体是否存在相关，也可以分别研究相关的方向，即是正相关，还是负相关。针对研究问题的不同，可以建立不同的假设组。

双侧检验

H_0:　不相关

H_1:　存在相关

单侧检验

H_0:　不相关　　H_0:不相关

H_+:　正相关　　H_-:负相关

为对假设作出判定，所需数据至少是定序尺度测量的。根据(7.4)式或(7.5)式、(7.6)式计算出 R 值。当 $n \leqslant 30$ 时，在附表 XⅢ 中，依据 n 和 R 查找相应的概率 P。这是 H_0 为真时，R 为某值可能的概率。若 P 值小于显著性水平 α，则数据拒绝 H_0；若 P 值大于显著性水平 α，则数据不能拒绝 H_0。表 7—8 是判定指导表。$n \leqslant 10$

表 7—8　　R 显著性检验判定指导表

备择假设	P 值(附表 XⅢ)
H_+:　正相关	R 的右尾概率
H_-:　负相关	R 的左尾概率
H_1:　存在相关	R 的较小概率的 2 倍

时，在附表 XⅢ 的第一部分查找，$10 < n \leqslant 30$ 时在表的第二部分查找相应的 P 值。若 $n > 30$，则按(7.7)式计算 Z。Z 统计量近似服从正态分布，可在附表 Ⅳ 中查找相应的 P 值。

$$Z = R\sqrt{n-1} \tag{7.7}$$

【例 7.4】　对例 7.3 作显著性检验

分析：由于例 7.3 中未指明相关的方向，只需检验是否相关，因而建立双侧备择：

H_0:　不相关

H_1:　存在相关

利用提供的数据计算的 R 值为 0.8491，每个样本数据 n 为 24。在附表 XⅢ 中，$n = 24$ 时，双侧检验 $|R| = 0.608$ 的概率为 0.002。显然，$R = 0.8491 > |R| = 0.608$。因此，概率 $P < 0.002$，数据拒绝 H_0，表明经济水平和卫生水平确实存在相关关系。

【例 7.5】 对例 7.2 进行显著性检验

分析：在例 7.2 中，结论指出经济水平与卫生水平存在正相关关系，因此，检验应是单侧备择：

H_0： 不相关

H_1： 正相关

根据表 7—5 计算的 $R = 0.8881$，在附表 XⅢ 中，$n = 12$ 时，$R = 0.825$ 的右尾概率是 0.001。显然，$R = 0.8881$ 的右尾概率 $P < 0.001$，数据不支持 H_0，表明该地区的经济发展水平和卫生水平存在正相关。

§7.2 Kendall 秩相关

Kendall 秩相关即肯德尔秩相关，与等级相关一样，也是用于两个样本相关程度的测量，要求数据至少是定序尺度的。它也是利用两组秩次测定两个样本间相关程度的一种非参数统计方法。

一、基本方法

n 个配对数据 $(x_1, y_1)(x_2, y_2), \cdots, (x_n, y_n)$ 分别抽选自 X、Y，X、Y 都至少是可以用定序尺度测量的。将 X 的 n 个数据的秩按自然顺序排列，则 Y 的 n 个秩也相应地发生变动。例如，X、Y 的秩分别为

X： 2 4 3 5 1

Y： 3 4 1 5 2

将 X 的秩按自然顺序排列后，X、Y 的秩则为下面的形式

X： 1 2 3 4 5

Y： 2 3 1 4 5

由于 X 的秩次已经按自然顺序由小到大排列，因此，X 的观察值每两个之间都是一致对。考察 Y 的秩次情况，第一个秩为 2，第二个为 3，因为 2 小于 3，是按自然顺序增加，因此，这是一个一致对。再考察 2 和 1，因为 2 大于 1，不是按自然顺序增加排列，所以这是一

个非一致对。依次考察下去，凡一致对记作 +1，非一致对记作 −1。考察结果如表 7—9 所示。

表 7—9　　Y 的数对评分

Y 的数对	分数	总和
2,3	1	8 个 +
2,1	−1	2 个 −
2,4	1	
2,5	1	
3,1	−1	
3,4	1	
3,5	1	
1,4	1	
1,5	1	
4,5	1	

在 X 的秩评定完全按自然顺序排列时，Y 的秩对所能给予的最大的评分，应是也完全按自然顺序排列的秩对的评分，即每一数对的评分均为 +1。这样，在 X、Y 的评秩完全一致的情况下，最大可能的评分总数应是一个组合，如在上例中是 $\binom{5}{2}=10$，即从 5 个里选 2 个的组合数。一般情况，n 个观察值对两两秩对之间评分，最大可能的总分为 $\binom{n}{2}$。以实际的评分与最大可能总分相比，可以测定两组秩之间的相关程度。

若以 U 表示 Y 的一致对数目，V 表示 Y 的非一致对数目，则一致对评分与最大可能总分之比为

$$\frac{U}{\binom{n}{2}}=\frac{2U}{n(n-1)} \tag{7.8}$$

非一致对评分与最大可能总分之比为

$$\frac{V}{\binom{n}{2}}=\frac{2V}{n(n-1)} \tag{7.9}$$

当 Y 的秩对完全按自然顺序排列时，(7.8) 式的值为 1，(7.9) 式的值为 0；而当 Y 的秩对全部为非一致对时，(7.9) 式的值为 1，(7.8) 式的值为 0。为测定两组秩之间的相关程度，定义的相关系数从 -1 到 $+1$，因此，Kendall 秩相关系数为

$$T=\frac{4U}{n(n-1)}-1 \tag{7.10}$$

$$T=1-\frac{4V}{n(n-1)} \tag{7.11}$$

若记 $S=U-V$，则 Kendall 秩相关系数为

$$T=\frac{2S}{n(n-1)} \tag{7.12}$$

这里的 Kendall 秩相关系数 T 是 Tau 的缩写，也常写作 τ。$T=1$，表明两组秩次完全正相关；$T=-1$，表明两组秩次间完全负相关。一般 $|T|>0.8$，可以为相关程度较高。

二、应用

【例 7.6】 利用例 7.2 的数据资料分析经济水平和卫生水平的相关程度

表 7—10　经济水平和卫生水平秩的排序

街道编号	经济水平的秩	卫生水平的秩	街道编号	经济水平的秩	卫生水平的秩
3	1	2	7	7	7
12	2	3	11	8	8
5	3	1	2	9	6
8	4	5	6	10	11
9	5	4	10	11	12
1	6	9	4	12	10

分析：根据表 7—5 的评秩结果进行秩次重新排列，将经济水平的

秩次按自然顺序排列，得表7—10的结果。由表可以计算得到2、3、1、5、4、9、7、8、6、11的一致对数目U

$$U = 10 + 9 + 9 + 7 + 7 + 3 + 4 + 3 + 3 + 1 = 56$$

根据(7.10)式计算有

$$T = \frac{4(56)}{(12)(12-1)} - 1 \doteq 0.6970$$

由表7—10可以计算出非一致对的数目V为

$$V = 1 + 1 + 1 + 3 + 1 + 1 + 1 + 1 = 10$$

根据(7.11)式

$$T = 1 - \frac{4(10)}{12(12-1)} = 1 - 0.3030 = 0.6970$$

若按(7.12)式计算，有

$$T = \frac{2S}{n(n-1)} = \frac{2(56-10)}{12(12-1)} = 0.6970$$

$T = 0.6970 < 0.8$，表明经济水平与卫生水平相关程度不够高。

Kendall秩相关系数也可以用于定距尺度测量的数据，数据不必评秩，而直接比较大小得到一致对或非一致对的数目。

【例7.7】 利用例7.2中表7—4的资料分析经济发展水平和卫生水平的相关程度

分析：将表7—4中经济发展水平的评分按从小到大的顺序排列，得到表7—11。由表计算卫生水平的一致对数目

$$U = 10 + 9 + 9 + 7 + 7 + 3 + 4 + 3 + 3 + 1 = 56$$

$$T = \frac{S}{\sqrt{n(n-1)/2 - u'}\ \sqrt{n(n-1)/2 - v'}} \quad (7.14)$$

【例 7.8】 两名裁判员对参赛 5 名歌手评秩的相关分析

两名裁判员对 5 名参赛歌手所评的等级如下

X 的秩	1	2.5	2.5	4.5	4.5
Y 的秩	2	3.5	3.5	1	5。

分析：由于 X、Y 中均有同秩，因此，X 的秩按自然顺序排列

表 7-12　　两个裁判员的秩的评分

X 的数对	Y 的数对	分数	总和
1,2.5	2,3.5	1	5 个加
1,2.5	2,3.5	1	3 个减
1,4.5	2,1	−1	
1,4.5	2,5	1	
2.5,2.5	3.5,3,5	0	
2.5,4.5	3.5,1	−1	
2.5,4.5	3.5,5	1	
2.5,4.5	3.5,1	−1	
2.5,4.5	3.5,5	1	
4.5,4.5	1,5	0	

时，同秩的两个值，可能对应 Y 的不同秩。这样应将所有可能考察的情况全部列出。评分结果如表 7—12。由表可知，无论是 X 或 Y 中，出现的数对是同秩时，其分数均为 0。根据表 7—12 的数据，$u' = \Sigma\binom{u}{2} = (1/2\Sigma u(u-1) = 1/2(2+2) = 2$，$v' = \Sigma\binom{v}{2} = 1/2\Sigma v(v-1) = 1$，于是由(7.14)式可得

$$T = \frac{5-3}{\sqrt{(1/2)5(5-1)-2}\ \sqrt{(1/2)5(5-1)-1}} \doteq 0.2357$$

由于 $T = 0.2357 < 0.8$，故两名裁判员对 5 名参赛歌手评秩

的相关程度极低。

对于 Kendall 秩相关系数来说，同分的影响也不是很大。

四、T 的显著性检验

与 Spearman 秩相关系数 R 一样，Kensall 秩相关系数 T 的显著性也应进行检验。这一检验实际上是检验两个总体的相关是否真实存在，是正相关或是负相关，从而说明以 T 的大小反映相关程度的高低是否可信。如果研究关心的是相关是否确实存在，而不考虑相关的方向，则应建立双侧备择，假设组为

H_0：　不相关

H_1：　存在相关

若关心的是相关的方向，则应建立单侧备择，假设组为

H_0：　不相关　　　H_0：　不相关

H_+：　正相关　　　H_-：　负相关

为对假设作出判定，所需数据至少是定序尺度测量的。通过对数据求出一致对或非一致对数目，按(7.10)式，(7.11)式或(7.14)式计算出 Kendall 秩相关系数 T。

T 的抽样分布在附表 XⅣ 中给出。当 $n \leqslant 10$ 时，只要根据 n、T 的值，可以在表的第一部分查找到 H_0 为真时，T 为某一值的概率 P；$10 < n \leqslant 30$ 时，在表的第二部分查找相应的概率 P。表 7—13 是判定指导表。若 $n > 30$，则按(7.15)计算 Z。

表 7—13　　T 显著性检验判定指导表

备择假设	P 值　(附表 XⅣ)
H_+：　正相关	T 的右尾概率
H_-：　负相关	T 的左尾概率
H_1：　存在相关	T 的较小概率的 2 倍

$$Z = \frac{3T\sqrt{n(n-1)}}{\sqrt{2(2n+5)}} = \frac{S}{\sqrt{n(n-1)(2n+5)/18}} \qquad (7.15)$$

由于 Z 近似正态分布，故可以在附表 Ⅳ 中查找相应的概率。

【例 7.8】 双胞胎智力的相关分析

某研究所对 10 对双胞胎儿童的智力进行调查，结果如表 7—14。

表 7—14　　儿童智力测试得分

双胞胎编号	先出生儿童(X)	后出生儿童(Y)
1	9.0	7.8
2	16.6	19.3
3	16.2	20.1
4	11.3	7.1
5	16.2	13.0
6	7.1	4.8
7	7.8	8.9
8	4.0	7.4
9	11.2	10.0
10	1.3	1.5

表 7—15　　儿童智力测试得分评秩

X	Y	X 的秩	Y 的秩	D	D^2
1.3	1.5	1	1	0	0
4.0	7.4	2	4	−2	4
7.1	4.8	3	2	1	1
7.8	8.9	4	6	−2	4
9.0	7.8	5	5	0	0
11.2	10.0	6	7	−1	1
11.3	7.1	7	3	4	16
16.2	20.1	8.5	10	−1.5	2.25
16.2	13.0	8.5	8	0.5	0.25
16.6	19.3	10	9	1	1

分析：智力测试得分是定距尺度测量的数据，可以转换为定序尺度。若将先出生儿童的得分记作 X，后出生儿童得分记作 Y，则将表 7—14 的得分评秩，得到表 7—15。利用这些数据分别计算

Spearman 秩相关系数 R 和 Kendall 秩相关系数 T。

由表 7—15 可以得到

$$R = 1 - \frac{6\Sigma D_i^2}{n(n^2 - 1)}$$

$$= 1 - \frac{6(29.5)}{10(10^2 - 1)}$$

$$\doteq 0.8212$$

由于 X 中有同分，故按(7.6)式计算，得到

$$R = \frac{10(10^2 - 1) - 6(29.5) - 1/2(3^2 - 2)}{\sqrt{10(10^2 - 1) - (2^3 - 2)}\sqrt{10(10^2 - 1) - 0}}$$

$$\doteq 0.8207$$

由表 7—15 可以计算得到 Y 秩的一致对数目

$$U = 9 + 6 + 7 + 4 + 4 + 3 + 3 + 0 + 1 = 37$$

非一致对数目

$$V = 0 + 2 + 0 + 2 + 1 + 1 + 0 + 1 + 0 + 0 = 7$$

Kendall 秩相关系数 T 为

$$T = \frac{2(U - V)}{n(n - 1)} = \frac{2(37 - 7)}{10(10 - 1)}$$

$$\doteq 0.6667$$

由于有同分，按(7.14)式计算得到

$$T = \frac{(37 - 7)}{\sqrt{(1/2)10(10 - 1) - (1/2)2(2 - 1)}\sqrt{(1/2)10(10 - 1)}}$$

$$\doteq 0.6742$$

对 R 和 T 的显著性进行检验，建立假设组

H_0: 不相关

H_+: 正相关

在附表 XⅢ 中，$n = 10$，$R = 0.8212$ 或 $R = 0.8207$ 相应的概率 $P < 0.003$；在附表 XⅣ 中，$n = 10$，$T = 0.6667$ 或 $T = 0.6742$ 相应的概率在 0.002 至 0.005 之间。这都表明数据不支持 H_0，也就是说双胞胎儿童的智力之间存在着正相关。按 Spearman 秩相关系数看，双胞胎儿童的智力之间相关程度较高。但按 Kensall 秩相关

反映两个样本间存在真正的或直接的关系。这种相关性是因为两个样本都和第三个样本有关系而产生。这个问题在参数统计中是通过偏相关解决的。在非参统计中也可以用偏相关方法处理。本节介绍的是 Kendall 偏秩相关系数 $T_{xy,z}$

一、基本方法

若 X、Y 与第三个样本 Z 有关，也就是说由于 Z 的变化对 X、Y 之间的关系有影响，那么，考察去掉 Z 的影响，仅仅研究 X、Y 之间的相关，就是偏相关。在统计上，偏相关就是在第三个样本 Z 保持恒定情况下，X、Y 之间的相关。

若有三个样本 X、Y、Z，每个样本均有 n 个数据，且都至少是在定序尺度上测量，那么根据Kendall 秩相关系数，T_{xy} 表示 X 与 Y 之间的秩相关程度，T_{xz} 表示 X、Z 间的秩相关程度，$T_{y,z}$ 则表示 Y 与 Z 间的秩相关程度。Kendall 偏秩相关系数 $T_{xy,z}$ 为

$$T_{xy,z} = \frac{T_{xy} - T_{xz}T_{yz}}{\sqrt{(1 - T_{xz}^2)(1 - T_{yz}^2)}} \tag{7.16}$$

$T_{xy,z}$ 是 Z 不变时，X 和 Y 之间的相关系数，有时也写作 $\tau_{xy,z}$。这里 T 仍是 Tau 的缩写。$T_{xy,z}$ 的取值范围也是从 -1 到 $+1$。但它的抽样分布至今未知。因而无法对其进行显著性检验。

二、应用

【例 7.9】 收入和继续受教育程度的相关分析

普遍认为在取得学士学位以后，在工作中继续研究生课程的学习很重要，也就是说在职继续学习比单纯地学院式学习更有意义。调查了 7 个人的年龄，花在继续学习上的时间以及月收入，结果如表 7—16。

表 7—16　　年龄、学习时间及收入

人员编号	年龄(岁)	继续学习的时间(小时)	月收入(千元)
1	40	18	2.5
2	35	0	2.0
3	30	6	1.5
4	36	15	2.4
5	41	24	3.0
6	45	30	2.8
7	48	45	2.9

分析：对于要分析研究的收入与继续学习的关系，在调查中增加了年龄，这可以看作是第三个样本。分析收入与继续学习的关系，可以利用表 7—16 的资料计算相关系数。虽然表中数据均为定比尺度测量，但样本数目较少，不符合参数统计中计算积矩相关系数的要求，且对总体不作任何假设，故应采用非参数统计方法。运用 Kendall 秩相关系数是最合适的。定义继续学习的时间为 X，月收入为 Y，年龄则为 Z。那么将 X 按从小到大的顺序排列后，Y 也可以得到一个新的排列如下

X:	0	6	15	18	24	30	45
Y:	2.0	1.5	2.4	2.5	3.0	2.8	2.9

$$U_{xy} = 5 + 5 + 4 + 3 + 0 + 1 + 0 = 18$$

则 $$T_{xy} = \frac{4(18)}{7(7-1)} - 1 = 0.7143$$

根据 $n = 7, T = 0.7143$，在附表 XIV 中查找相应的概率 P 近似于 0.015，对于显著性水平 $\alpha = 0.05$，P 值足够小，因此数据拒绝不相关的零假设，表明收入与继续学习的时间存在相关。由于 $T = 0.7143$，表明二者相关程度不算很高。

但考虑到收入可能受到年龄的影响，继续学习的时间也会受到年龄的影响，因而应排除年龄因素的影响，计算偏相关系数。对于 Kendall 秩相关系数来说，最易推广的偏相关系数就是偏秩相关系数 $T_{xy,z}$。根据(7.16)式的计算公式，需分别计算 T_{xz}，T_{yz}。

将 X 依次排序，得到 Z 的排列如下：

X:	0	6	15	18	14	30	45
Z:	35	30	36	40	41	45	48

$$U_{xz} = 5 + 5 + 4 + 3 + 2 + 1 + 0 = 20$$

$$T_{xz} = \frac{40(20)}{7(7-1)} - 1 = 0.9048$$

将 Y 依次排序，得到 Z 的排列为

Y:	1.5	2.0	2.4	2.5	2.8	2.9	3.0
Z:	30	35	36	40	45	48	41

$$U_{yz} = 6 + 5 + 4 + 3 + 1 + 0 + 0 = 19$$

$$T_{yz} = \frac{4(19)}{7(7-1)} - 1 = 0.8095$$

将上述结果代入(7.16)式，得到

$$T_{xy,z} = \frac{0.7143 - (0.9048)(0.8095)}{\sqrt{1 - 0.9048^2}\ \sqrt{1 - 0.8095^2}} = -0.0724$$

$T_{xy,z}$ 与 T_{xy} 比较起来相差甚远。$T_{xy,z}$ 仅仅是一个很小的数值，以至无法作出存在相关的结论。但它与 T_{xy} 共同使用，可以说明，收入与继续学习的时间，在年龄因素的影响下存在一定的相关。除了年龄的影响外，二者之间几乎没有什么联系。

第八章　k 个样本的相关分析

前面一章研究的是 n 个对象或个体的两组秩之间相关的度量，在实际问题中，往往还涉及 n 个对象或个体的几组秩评定之间的相关。对于至少是定序尺度测量的 k 个配对样本的数据，或 k 次试验得到的数据，其秩评定间的相关，可以采用 Kendall 秩评定协和系数度量。本章主要介绍两种 Kendall 秩评定协和系数：完全秩评定协和系数和不完全秩评定协和系数。

§8.1　完全秩评定的 Kendall 协和系数

完全秩评定的 Kendall 协和系数(Kendall Coefficient of Concordance for Complete Rankings)用于 k 组秩评定间相关程度的测定，即多组秩之间关联程度的测定。

一、基本方法

若被分析的数据是定序尺度测量的，那么 n 个数据，即 n 个对象或个体，可以分别给予某一个秩，在这一组数据内所有的秩次和即等级和为

$$1 + 2 + 3 + \cdots + n = n(n + 1)/2$$

如果有 k 组秩，那么这 k 组秩的秩次总和就是 $kn(n + 1)/2$。

例如 3 个消费者分别给 6 种牌号电冰箱的质量评等级，结果如表 8—1。表中最后一行 R_j 是每一种牌号电冰箱的秩和，总的秩和为 $kn(n + 1)/2 = 3(6)(6 + 1)/2 = 63 = \sum_{j=1}^{6} R_j$。这也就是最大可能的秩次和。这时，对于每一个观察对象或个体来说，平均的秩次和应为

$$[kn(n + 1))/2]/n$$

表 8—1　　消费者给电冰箱质量评定的秩

消费者	冰箱 A	冰箱 B	冰箱 C	冰箱 D	冰箱 E	冰箱 F
1	1	6	3	2	5	4
2	1	5	6	4	2	3
3	6	3	2	5	4	1
秩和(R_j)	8	14	11	11	11	8

即为 $k(n+1)/2$。如果 $R_j(j=1,2,\cdots,n)$ 表示每一观察对象或个体的实际秩和，那么，R_j 与 $k(n+1)/2$ 越接近，表明对第 j 个观察对象或个体的秩评定越接近平均秩；二者相差越远，远离平均秩。由于 R_j 与 $k(n+1)/2$ 的差值可正可负，因此，在分析时应采用差值的平方和。定义差值的平方和为 S，即

$$S=\sum_{j=1}^{n}[R_j-k(n+1)/2]^2 \tag{8.1}$$

在 k 组秩评定完全一致时，各个观察对象或个体的秩和与平均秩和的离差平方和，是最大可能的离差平方和。由于 k 组秩评定完全一致时，各观察对象或个体的秩和分别为 $k,2k,\cdots,nk$，如表 8—1，如果 3 位消费者对 6 种牌号电冰箱的质量看法一致，那么他们会给出相同的秩。这时，被认为质量最好的电冰箱将得到 3 个秩 1，它的秩和 $R_j=1+1+1=3=k$。被认为质量第二的，秩和 $R_j=2+2+2=6=2k$。最差的电冰箱秩和将是 $R_j=6+6+6=18=nk$。也就是说，当 k 组秩评定之间完全一致的时候，R_j 应是 k，$2k,\cdots,nk$。因此，最大可能的离差平方和为

$$\sum_{j=1}^{n}[jk-k(n+1)/2]^2=k^2\sum_{j=1}^{n}[j-(n+1)/2]^2$$
$$=k^2n(n^2-1)/12 \tag{8.2}$$

实际偏差平方和与最大可能偏差平方和之比，在一定程度上能反映 k 组秩评定间的一致性，即协调程度。(8.1) 式除以 (8.2) 式得到 Kendall 完全秩评定协和系数 W。

$$W=\frac{12S}{k^2n(n^2-1)}$$

$$= 12\sum_{j=1}^{n}\frac{[R_j - k(n+1)/2]^2}{k^2n(n^2-1)} \qquad (8.3)$$

W 的取值在 0 到 1 之间。若 $W = 0$,表明 k 组秩之间不相关;若 $W = 1$,表明 k 组秩之间完全相关,即完全一致。由于 $k > 2$ 时,k 组秩评定不可能完全不一致,也就是说,只有当 $k = 2$ 时,秩评定一致和非一致是对称相反的,而 $k > 2$,对称性不再存在,因此,W 取值不可能为负。

为方便实际计算,(8.3) 式还可以写成下面的形式

$$W = \frac{12\sum_{j=1}^{n} R_j^2 - 3k^2n(n+1)^2}{k^2n(n^2-1)} \qquad (8.4)$$

二、应用

【例 8.1】 裁判组整体评分效果的相关分析

在某次业余歌手大赛上,6 名裁判员组成的裁判组,对 10 名参赛歌手的评分等级如表 8—2。

表 8—2　　　　裁判员对歌手评定的等级

歌手编号	裁判员 A	裁判员 B	裁判员 C	裁判员 D	裁判员 E	裁判员 F
1	1	1	1	1	1	5
2	2	3	7	5	8	3
3	5	4	6	2	6	9
4	9	7	5	4	10	6
5	4	5	3	6	5	8
6	6	6	4	7	7	2
7	3	2	9	10	2	4
8	7	10	10	3	4	1
9	10	8	2	9	9	7
10	8	9	8	8	3	10

分析:裁判组由 6 名裁判组成,要评价 6 名裁判整体评分效果,实际上是评价裁判组整体评分的一致程度。由于这是 6 个配对

出现同分时，(8.2)式就变成

$$\frac{k^2n(n^2-1)-k(\Sigma t^3-\Sigma t)}{12}$$

式中的t是同分的观察值数目。这样，(8.3)式和(8.4)式的分母就应该是

$$k^2n(n^2-1)-k(\Sigma t^3-\Sigma t) \qquad (8.5)$$

经过同分校正后的Kendall完全秩评定协和系数W为

$$W=\frac{12\sum_{j=1}^{n}R_j^2-3k^2n(n+1)^2}{k^2n(n^2-1)-k(\Sigma t^3-\Sigma t)} \qquad (8.6)$$

同分会使(8.4)式计算的W值偏低。当同分观察值的数目所占比例较小时，这种影响能够被忽略，仍用(8.4)式计算W；但当同分观察值的数目较多，所占比重较大时，这种影响不应被忽略，应采用(8.6)式计算W。

【例8.2】 裁判组整体评分效果的相关分析

某次业余歌手大赛，6名裁判员对10名参赛歌手成绩评分后，转换为秩，如表8—4。

表8—4　　裁判员对歌手的评分等级

歌手编号	裁判员A	裁判员B	裁判员C	裁判员D	裁判员E	裁判员F
1	5.5	1	1	1	1	1.5
2	2.5	3	4	2	8	1.5
3	5.5	7.5	2.5	3.5	6	7.5
4	4	7.5	2.5	3.5	10	9.5
5	7	5	6	9	5	3.5
6	8	2	8.5	6	7	6
7	1	10	8.5	8	2	9.5
8	9	9	6	5	4	3.5
9	2.5	4	6	10	9	5
10	10	6	10	7	3	7.5

分析：这一问题与例8.1类似，也应采用完全秩评定的

Kendall 协和系数度量裁判组整体评分效果的一致性。由于有较多的同分，因而必须进行校正。表8—5是R_j的计算表。在裁判员A给的10名歌手评分中，有2个2.5秩，2个5.5秩，因而在裁判员A的评秩中Σt^3-$\Sigma t = (2^3 + 2^3) - (2 + 2)$；在裁判员$B$所评定的秩中，同分的只有7.5，裁判员$c$有2个2.5，2个8.5，3个6，以此类推，并计算$\Sigma t^3 - \Sigma t$。计算过程如表8—6。

表8—5　　R_j的计算表

歌手编号	R_j	R_j^2
1	11	121
2	21	441
3	30.5	930.25
4	37	1369
5	35.5	1260.25
6	37.5	1406.25
7	39	1521
8	36.5	1332.25
9	36.5	1332.25
10	43.5	1892.25
合计	328	11605.5

表8—6　　同分校正因子计算表

	t	t^3
A	$2+2$	2^3+2^3
B	2	2^3
C	$2+2+3$	$2^3+2^3+3^3$
D	2	2^3
F	$2+2+2+2$	$2^3+2^3+2^3+2^3$
合计	23	107

将$n = 10$，$k = 6$以及表8—5，表8—6的计算结果代入(8.6)式，

得到校正后的 W 为

$$W = \frac{12(11605.5) - 3(6^2)10(10+1)^2}{(6^2)10(10^2-1) - 6(107-23)}$$

$$\doteq 0.2444$$

若不进行校正，直接运用(8.4)式计算，得到

$$W = \frac{12(11605.5) - 3(6^2)10(10+1)^2}{10(6^2)(10^2-1)}$$

$$\doteq 0.2409$$

计算结果表明，6 名裁判员的裁判组对参赛歌手的看法一致程度很低。校正前的 W 值低于校正后的 W 值。因此，当同分的观察值数目比例较大时，一定要进行校正。

四、W 的显著性检验

对 W 显著性的检验，是为了对总体间是否存在真实的相关关系作出判定。由于是 k 个样本，只能建立双侧备择，假设组为

H_0：　不相关

H_1：　存在相关

为了对假设作出判定，需要容量均为 n 的 k 个样本数据至少是在定序尺度上测量的，每一观察值都能有相应的秩。

检验统计量因样本的大小而有所不同。当样本的观察值 n 较小时，采用的检验统计量为 S。

$$S = \sum_{j=1}^{n} R_j^2 - 3k^2n(n-1)^2 \tag{8.7}$$

当样本观察值数目 n 较大时，采用 Q 统计量。

$$Q = k(n-1)W \tag{8.8}$$

或

$$Q = \frac{12S}{kn(n+1)} \tag{8.9}$$

检验统计量 S 在 H_0 为真时的抽样分布，如附表 XV。当 $n=3, k \leqslant 8$ 时，与 S 值相对应的概率 P 值，以及 $n=4, k \leqslant 4$ 时，S 值出现的概率 P 值，均可以在附表 XV 中查找。其它时候，利用(8.8)式或(8.9)式计算 Q 统计量。Q 统计量近似为自由度 $\mathrm{df} = n-1$ 的

χ^2 分布,在附表 I 中可以查找 H_0 为真时,Q 为某值的概率。无论是在附表 XV 或附表 I 中查找的概率 P 值,若小于或等于给定的显著性水平 α,则数据拒绝 H_0,表明总体存在相关。

【例 8.3】 对例 8.2 的 W 值进行显著性检验

分析:因为在这个问题中,$n = 10$,$k = 6$,所以应利用 W 值计算统计量 Q,根据(8.8) 式计算得到

$$
\begin{aligned}
Q &= k(n-1)W \\
&= 6(10-1)(0.2444) \\
&= 13.1976
\end{aligned}
$$

根据自由度 $\mathrm{df} = n - 1 = 9$,显著性水平 $\alpha = 0.05$,在附表 I 中查找得到 $\chi^2_\alpha = 16.92$。由于

$$Q = 13.1976 < \chi^2_\alpha = 16.92$$

所以数据在 5% 的水平上不能拒绝 H_0,表明 6 名裁判员所作的秩评定彼此不相关。若在附表 I 中,查找自由度 $\mathrm{df} = n - 1 = 9$,在 H_0 为真时,$Q = 13.1976$,出现的概率 P 在 0.10 到 0.20 之间。这一概率显然大于显著性水平 $\alpha = 0.05$,因此可以作出结论:在 H_0 为真时,W 值出现的概率较大,不能拒绝 H_0,也就是说 6 名裁判员对参赛歌手所作的秩评定彼此无关。

§8.2 不完全秩评定的 Kendall 协和系数

在实际问题中,往往会遇到这样的情况,如在参赛的 10 名歌手中,只评出 6 名排等级;在对几种消费品质量评级时,消费者只评出其中最满意的 3 种等等。这时,不是所有配对样本的每个观察值都被分配等级,也就是秩的评定不完全。研究这种情况下,裁判员评分效果的一致程度,消费者对产品质量满意的一致程度等,不能采用上节所述的 Kendall 协和系数,而应采用本节介绍的不完全秩评定的 Kendall 协和系数。

一、基本方法

若被分析的数据是 k 个组,即 k 个样本,每组均含有 n 个观察

值，对每组观察值评定的秩不是n个而是m个，且$m<n$，则构成不完全秩评定的情况，可以考虑采用不完全秩评定的 Kendall 协和系数。但在使用时，通常有这样的限制，即对于m,n,k以及λ来说，应该是匹配的，即满足(8.10)式。

$$\lambda n(n-1)=km(m-1) \tag{8.10}$$

式中，λ是配对样本被比较的次数。例如，3 个消费者对 3 种牌号的彩电质量评等级，若 3 种牌号的彩电仅一次被比较，则$\lambda=1$。(8.10)式是不完全秩评定的 Kendall 协和系数运用的一个假设，称作平衡假设。m、n、k之间不是能够完全任意的。如对于 4 个观察对象的一个配对比较，$n=4$，$m=2$，当$\lambda=1$时，要求$k=6$；若$\lambda=2$，则$k=12$，也就是说，当有 4 名参赛歌手比赛时，按不完全秩评定的设计，从中评出 2 名给予等级，那么一次比较时，需要 6 名裁判员，若两次被比较，则需要 12 名裁判员。一般情况下，λ、m、n、k的取值如表 8—7。

当分析的数据符合上述条件时，可以得到一个k行和n列的表。在每一行中，只有m个秩，在每一列中，有km/n个秩。对于每一行来说，秩和为$1+2+\cdots+m=m(m-1)/2$，由于有k行，

表 8—7

λ	k	n	m
1	3	3	2
1	6	4	2
1	7	7	3
1	10	5	2
2	4	4	3
2	6	3	2
2	12	4	2

因此，所有的秩和为$km(m+1)/2$。对于n列来说，平均每列的秩和为

$$km(m+1)/2n$$

若每列的实际秩和记作 $R_j(j=1,2,\cdots,n)$ 那么，实际秩和与平均秩和的差值大小也可以用作判定协调性程度。由于实际秩和与平均秩和之差可正可负，因而，仍采用差值平方和。定义 S 为

$$S=\sum_{j=1}^{n}\left[R_j-\frac{km(m+1)}{2n}\right]^2 \tag{8.11}$$

当 k 组秩的评定完全一致时，S 就是

$$\frac{\lambda^2 n(n^2-1)}{12} \tag{8.12}$$

这也就是实际秩和与平均秩和差值平方和的最大可能值。因此，将(8.11)式与(8.12)式相比，可以用来度量 k 组秩评定之间的协调性或一致性。这就是不完全秩评定的 Kendall 协和系数，也记作 W。其计算公式如(8.13)式和(8.14)式。

$$W=\frac{12S}{\lambda^2 n(n^2-1)} \tag{8.13}$$

$$=12\sum_{j=1}^{n}\frac{[R_j-km(m+1)/2n]^2}{\lambda^2 n(n^2-1)} \tag{8.14}$$

如果 $m=n$ 和 $\lambda=k$，那么(8.13)式和(8.14)式就是(8.3)式。

W 值在0到1之间。W 为0，表明 k 组秩评定之间不相关；W 为1，表明 k 组秩评定之间完全相关，即完全一致。由于 $k>2$ 时，秩评定的一致和非一致不是对称的，因而，W 取值不可能负。为实际计算的方便，(8.13)式和(8.14)式可以写成(8.15)式的形式。

$$12S=12\Sigma R_j^2-\frac{3k^2m^2(m+1)^2}{n}$$

$$W=\frac{12\sum_{j=1}^{n}R_j^2-3k^2m^2(m+1)^2/n}{\lambda^2 n(n^2-1)} \tag{8.15}$$

二、应用

【例 8.4】 消费者对彩电质量评价的一致性分析

7种不同牌号的彩电质量检验，不要求消费者对每一种牌号的彩电都给出秩，只要求不大于3个。因而，$m=3,n=7$。如果每

对彩电仅一次被比较，因而有 $\lambda = 1$。需要的消费者数目，可以从(8.10)式中计算得到

$$k = \frac{\lambda n(n-1)}{m(m-1)} = \frac{1(7)(7-1)}{3(3-1)} = 7$$

表 8—8　　消费者对彩电质量可能的评秩

消费者编号	彩电 A	彩电 B	彩电 C	彩电 D	彩电 E	彩电 F	彩电 G
1	*	*		*			
2		*	*		*		
3			*	*		*	
4				*	*		*
5	*				*	*	
6		*				*	*
7	*		*				*

即需要有 7 名消费者来评定。表 8—8 是一种满足要求的设计表格。表中的 * 号表示那种牌号的彩电被消费者评秩。由于 $m = 3$，$n = 7$，$k = 7$，所以表中每一行有 $m = 3$ 个秩，表中每一列有 $km/n = 7(3)/7 = 3$ 个秩。经过消费者评秩，结果如表 8—9。

表 8—9　　消费者对彩电的秩评定结果

消费者编号	A	B	C	D	E	F	G
1	1	2		3			
2		1	3		2		
3			3	2		1	
4				2	3		1
5	1				3	2	
6		2				1	3
7	1		3				2
合计(R_j)	3	5	9	7	8	4	6

分析：由于这是不完全的秩评定，设计要求符合(8.10)式的平衡假定，可以采用不完全秩评定的 Kendall 协和系数，分析消费者对彩电质量评价的一致性。

根据(8.15)式有

$$
\begin{aligned}
12S &= 12(3^2 + 5^2 + 9^2 + 7^2 + 8^2 + 4^2 + 6^2) \\
&\quad - 3(7^2)(3^2)(3+1)^2/7 \\
&= 12(280) - 3024 \\
&= 336
\end{aligned}
$$

$$W = \frac{336}{7(7^2 - 1)} = 1$$

这是 W 的最大值，表明 7 个消费者对彩电质量的看法完全一致。

在不完全的秩评定中，同分也是可能出现的，因为数据可以由定距尺度的评分转换为定序尺度的秩。但是，目前没有比较简单的校正(8.12)式的公式，因此，计算 Kendall 协和系数时，仍旧采用(8.13)式、(8.14)式或(8.15)式。

三、显著性检验

对于不完全秩评定的 Kendall 协和系数，也可以进行显著性检验。建立的假设组为

H_0：　不相关

H_1：　存在相关

为对假设作出判定，需要 k 个样本的数据至少是定序尺度测量的，并能够根据(8.11)式、(8.15)式分别计算出 S、W。利用 S、W，按照(8.16)式、(8.17)式计算得到检验统计量 Q。统计量 Q 近似于自由度 $df = n - 1$ 的 χ^2 分布。因此，在附表 I 中，可以查找 H_0 为真时，Q 为某值，即 W 为某值的概率 P。根据 P 的大小，可以对拒绝或不能拒绝 H_0 作出判断。

$$Q = \frac{\lambda(n^2 - 1)W}{m + 1} \tag{8.16}$$

$$= \frac{12S}{\lambda n(m + 1)} \tag{8.17}$$

【例 8.5】 利用例 8.4 的数据作显著性检验

分析：在例 8.4 中，$\lambda = 1, m = 3, n = 7, W = 1$，将各个数值代入(8.16) 式得到

$$Q = \frac{1(7^2 - 1)1}{3 + 1} = 12$$

自由度 $\mathrm{df} = n - 1 = 6$，附表 Ⅰ 中，H_0 为真时，Q 是 12 出现的概率 P 略大于 0.05，因为当概率为 0.05 时，$\chi_a^2 = 12.59$。由于这个 P 是近似的值，因而，可以在显著性水平 $\alpha = 0.05$ 上拒绝 H_0。况且，W 的值为最大可能值 1，拒绝 H_0 是合乎逻辑的。

四、多重比较

若 P 值很小，以至于 H_0 被拒绝，也就是 k 组不完全的秩评定之间存在相关，那么，有必要比较这 k 组秩之间是否有重大的不同，可以运用多重比较技术来研究这一问题。

第 j 列的秩和 R_j 除以该列的秩数目 km/n，是第 j 个对象的平均秩，比较 n 个对象的 k 个不完全组的秩之间有无明显不同，可以利用各个对象的平均秩比较，也可以利用各列秩和的差值比较。对于任意两个列等级和$(R_i, R_j)1 \leqslant i \neq j \leqslant n$，它们的差值在 $1 - \alpha$ 水平下，满足(8.18) 式或(8.19) 式

$$|R_i - R_j| \leqslant Z\sqrt{\frac{km(m^2 - 1)}{6(n - 1)}} \tag{8.18}$$

或

$$|R_i - R_j| \leqslant Z\sqrt{\frac{n\lambda(m + 1)}{6}} \tag{8.19}$$

表明 n 个对象间秩评定没有重大的不同；而若 $|R_i - R_j|$ 大于(8.18) 式或(8.19) 式的右侧，表明这是有重大不同的数对。

(8.18) 式、(8.19) 式中的 Z，当 n 较大时，在附表 Ⅳ 中，利用 $\alpha/n(n - 1)$ 查找。因为 Z 是对应于 $\alpha/n(n - 1)$ 右尾概率的正态曲

线的临界值点。若 n 不大，在附表 XⅥ 中以 $P=n(n-1)/2$，显著性水平为 α，可以查找到相应的数值，替代(8.18) 式、(8.19) 式中的 Z。

【例 8.6】 对例 8.4 的各组秩进行比较

分析：在例 8.5 中，由于 P 值不够大，拒绝了 H_0，即各组不完全秩评定间没有差别的零假设被拒绝。在这个前提下，可以进行多重比较。由表 8—9 所得到的各列秩和可知，消费者对 7 种牌号彩电质量的评价最好的为 A，最差的为 C。根据秩和由少到多排列，消费者对 7 种彩电质量满意程度依次为：A、F、B、G、D、E、C。

取显著性水干 $\alpha=0.10$，由于 $n=7$，可得 $P=n(n-1)/2=21$，查附表 XⅥ，得到 2.823，代入(8.19) 式，得到

$$|R_i-R_j|\leqslant 2.823\sqrt{7(1)(3+1)/6}\doteq 6.0984$$

将表 8—9 中各列秩和两俩比较差值，可知，最大差值为 A 与 C 列即 $|3-9|=6$。这一差值仍小于 6.0984。因此，可以得出结论：在 7 种牌号彩电的不完全秩评定中，各列之间没有什么重大不同，若取 $\alpha=0.30$，由(8.19) 式得到

$$|R_i-R_j|\leqslant 2.45\sqrt{7(1)(3+1)/6}\doteq 5.2926$$

根据表 8—9 中提供的各列秩和可知，除了 A 与 C 秩和的差值较大外，其它各列秩和之间没有什么重大的不同，A 与 C 的差异，说明对于消费者来说，对彩电 A 质量的评价大大高于对彩电 C 的评价，因为 A 的秩次和低于 C 的秩次和较多。

§8.3 Friedman 检验和多重比较

Kendall 协和系数是 k 个样本的秩评定间一致性的测度，可用于检验 k 组秩之间的独立性。有时，研究所关心的并不是 k 组秩之间是否独立，而是 k 组秩之间的联系。例如，对同一种疾病，有几种处理，那么几种处理效果是否相同，对这样的问题进行检验并作多重比较，可以采用 Friedman 检验。

Friedman 检验亦称弗里德曼双向评秩方差分析(Friedman's

Two-way analysis of variance by ranks 或 Friedman Two-way analysis with ranks）它是在第五章中所介绍的方法。

一、基本方法

若研究的是几种处理的效果是否有差异，那么建立的假设组为：

H_0：$\mu_1 = \mu_2 = \cdots \mu_n$

H_1：$\mu_j (j = 1, 2, \cdots, n)$ 中至少有两个不等

这里 $\mu_j (j = 1, 2, \cdots, n)$ 是第 j 种处理的平均效果，即平均秩。为对假设作出判定，所需要的数据至少是在定序尺度上测量，要能将数据排成 k 行 n 列的双向表，如下面所示。表中 R_{ij} 表示第 i 行第 j 列数据所应评定的秩。对于每列来说，平均秩和应是 $[kn(n+1)/2]/n = k(n+1)/2$。对 H_0 是否成立的判定，可以通过各个 $R_j (j = 1, 2, \cdots, n)$ 与平均秩和 $k(n+1)/2$ 之差的大小来作出定义

$$
\begin{array}{lccccccc}
 & & & \text{处} & & \text{理} & & \text{行的秩和} \\
 & R_{11} & R_{12} & \cdots & R_{1j} & \cdots & R_{1n} & n(n+1)/2 \\
 & R_{21} & R_{22} & \cdots & R_{2j} & \cdots & R_{2n} & n(n+1)/2 \\
 & \vdots & & & \vdots & \vdots & & \vdots \\
\text{组} & R_{i1} & R_{i2} & \cdots & R_{ij} & \cdots & R_{in} & n(n+1)/2 \\
 & \vdots & & & \vdots & \vdots & & \vdots \\
 & R_{k1} & R_{k2} & \cdots & R_{kj} & \cdots & R_{kn} & n(n+1)/2 \\
\text{列的和秩} & R_1 & R_2 & & \cdots & & R_n & kn(n+1)/2
\end{array}
$$

$$S = \sum_{j=1}^{n} \left[R_j - \frac{k(n+1)}{2} \right]^2 \tag{8.20}$$

S 是实际的列秩和与平均秩和的离差平方和。这与（8.1）式是相同的。它也可以写成

$$S = \sum_{j=1}^{n} \left[\sum_{i=1}^{k} \left(R_{ij} - \frac{n+1}{2} \right) \right]^2 \tag{8.21}$$

为对假设作出判定，定义统计量 Q 为检验统计量

$$Q = \frac{12S}{kn(n+1)}$$

$$= 12\sum_{j=1}^{r} R_j^2 / kn(n+1) - 3k(n+1) \qquad (8.22)$$

这与(8.9)式相同。统计量Q近似为自由度$df = n-1$的χ^2分布，在附表Ⅰ中，可以查找到H_0为真时，Q为某值的概率P；也可以根据给定的显著性水平α，$df = n-1$在附表Ⅰ中查得H_0为真时的临界值χ_α^2，若$Q \geq \chi_\alpha^2$，则拒绝H_0；若$Q < \chi_\alpha^2$，则不能拒绝H_0。

二、应用

【例 8.7】 四种药物治疗效果是否有差异

四种药物可以治疗同一种疾病，为检验其效果是否有显著差异，选取了32名患该种病的病人接受治疗。每4人一组，共分8组。同一组的4名病人符合配对条件，随机地指定某人使用药物A、B、C，剩余1人用药物D。经过一个周期的治疗后，根据病情好转情况评分，结果如表8—10。

表 8—10 病情好转得分

病人组	药物 A	药物 B	药物 C	药物 D
1	14	23	26	30
2	19	25	25	33
3	17	22	29	28
4	17	21	28	27
5	16	24	28	32
6	15	26	27	26
7	18	26	27	36
8	16	22	30	32

分析：评价8个配对样本的差异性，实际是检验考察8组秩是否相同，可以利用各样本的平均秩是否相同，完成检验目的。建立的假设组为

$$H_0: \quad \mu_1 = \mu_2 = \mu_3 = \mu_4$$

H_1: $\mu_j(j=1,2,3,4)$ 中至少有两个不等

这是样本容量为 4 的 8 个相关样本的检验问题，可以采用 Friedman 检验。将表 8—10 中的评分转换为秩，并求出每列的秩和 R_j。如表 8—11 所示。

根据(8.22)式可以得到

$$Q = 12(8^2 + 17^2 + 26.5^2 + 28.5^2)/4(4+1) - 3(8)(4+1) = 22410/160 - 120 = 20.0625$$

查附表 I 中，显著性水平 $\alpha = 0.05$, $df = n - 1 = 3$, H_0 为真时的临界值 $\chi^2_\alpha = 7.82$。由于 $Q = 20.0652 > \chi^2_\alpha = 7.82$，所以数据不支持 H_0，即四种药物的治疗效果有显著差异。若以 $df = n - 1 = 3$，查找附表 I，$Q = 20.0625$ 出现的概率 $P < 0.001$。显然 P 值大大小于显著性水平 $\alpha = 0.05$。因此，调查结果不支持 H_0。

表 8—11　　病情好转程度的秩

病人的组	药物 A	药物 B	药物 C	药物 D
1	1	2	3	4
2	1	2.5	2.5	4
3	1	2	4	3
4	1	2	4	3
5	1	2	3	4
6	1	2.5	4	2.5
7	1	2	3	4
8	1	2	3	4
合计(R_j)	8	17	26.5	28.5

三、多重比较

若在上述检验中 P 值较小，拒绝了 H_0，也就是拒绝了几种处

理效果之间没有差异的零假设。这时，可以进一步研究哪种处理的效果与其它的处理有明显的不同。这就要进行多重比较。

在 $k \times n$ 表中，秩的列总和 R_j 被 k 除以后是一个平均秩；若以 $\mu_j(j=1,2,\cdots,n)$ 表示秩的平均值，则 R_j/k 是 μ_j 的一个估计值。多重比较就是利用 μ_j 之间的差异比较几种处理有无重大不同的一种方法。它是借助于一个全面或总的评价关系进行比较。使用平均秩 R_j/k 与使用秩列的总和 R_j 是等价的，因而与 §8.2 中一样，可以利用两个列秩和的差值进行比较。在 $1-\alpha$ 的水平下，下面的不等式(8.23)对于所有的列秩和对 $(R_i,R_j)1 \leqslant i \neq j \leqslant n$ 都能成立，则几种处理之间没有什么重大的不同。若 $|R_i-R_j|$ 大于右侧的值，则表明这两种处理之间有重大的不同。

$$|R_i - R_j| \leqslant Z\sqrt{\frac{kn(n+1)}{6}} \tag{8.23}$$

(8.23)式中的常数 Z 是正态曲线的一个临界值点，它对应于 $\alpha/n(n-1)$ 的右尾概率。借助于 n 和 α，可以在附表 Ⅳ 中查到相应的 Z 值。当 n 较小时，借助于 $P=n(n-1)/2$ 计算得到 P，在附表 XⅥ 中，查找与 α 相应的值，这就是(8.23)式中的 Z 值。

【例 8.8】 四种药物治疗效果之间有无重大不同

利用例 8.7 的数据比较四种药物治疗效果间有无重大不同

分析：例 8.7 中 H_0 被拒绝，表明四种药物对治疗同一种疾病的效果不完全相同。但它们之间有无重大差异，哪两个之间最不相同，可以运用多重比较加以解决。

由于 n 不大，因而计算 $P=n(n-1)/2=6$。附表 XⅥ 中，$\alpha=0.10$，$P=6$，Z 值为 2.394，代入(8.23)式得到

$$|R_i - R_j| \leqslant 2.394\sqrt{\frac{8(4)(4+1)}{6}} \doteq 12.3626$$

计算表8—11中各列秩和之差，可以得知，对于 $\alpha=0.10$ 的水平，A 药物和 C 药物，B 药物和 D 药物的秩和之差均大于 12.3626，表明它们的治疗效果之间有重大的不同。由于评秩是根据评分由小到

教育水平。表9—1所列出的两个变量之间，是考察教育水平对妇女志愿的影响。这种情况称为不对称关系。

列联表可以清楚地反映在X变量条件下，Y的次数分布情况。因此，列联表又称作条件次数表。表的最下端是每列的总次数，称为行边缘次数，表的最右列是每行的总次数，称为列边缘次数。表中的次数，称为条件次数，表示在自变量每个条件下，因变量各个值的数目。例如，在表9—1中，教育水平低的有200人，这是边缘次数，其中以理想工作为志愿的有105人，这是条件次数。

表9—1只有2行和2列，因此，称为2×2列联表。这是最简单的列联表。当行、列不只为2时，一般称为$r\times c$列联表。其中r(Row)表示行，c(Column)表示列。若$f_{ij}(i=1,2,\cdots,r;j=1,2,\cdots,C)$表示条件次数，则$r\times c$列联表如表9—2所示。

表9—2　　$r\times c$ 列联表

	X_1	X_2	$\cdots$	X_j	$\cdots$	X_C	合　计
Y_1	f_{11}	f_{12}	$\cdots$	f_{1j}	$\cdots$	f_{1C}	$f_{1\cdot}$
Y_2	f_{21}	f_{22}	$\cdots$	f_{2j}	$\cdots$	f_{2C}	$f_{2\cdot}$
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$	
Y_i	f_{i1}	f_{i2}	$\cdots$	f_{ij}	$\cdots$	f_{iC}	$f_{i\cdot}$
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$	
Y_r	f_{r1}	f_{r2}	$\cdots$	f_{rj}	$\cdots$	f_{rC}	$f_{r\cdot}$
合　计	$f_{\cdot 1}$	$f_{\cdot 2}$	$\cdots$	$f_{\cdot j}$	$\cdots$	$f_{\cdot C}$	n

$r\times c$列联表中，r和c可以不相等，即可以是3×3列联表，也可以3×2或2×3列联表。表9—2中，$f_{\cdot j}=\sum_{i=1}^{r}f_{ij}$，$f_{i\cdot}=\sum_{j=1}^{c}f_{ij}$，$n=\sum_{i=1}^{r}\sum_{j=1}^{c}f_{ij}=\sum_{i=1}^{r}f_{i\cdot}=\sum_{j=1}^{c}f_{\cdot j}$。

二、条件百分表

从条件次数表虽然可以知道在X的条件下，Y变量值的次数，

布不同于总体中的分布。这时，以自变量的方向计算百分数会歪曲资料，需要按因变量的方向计算。例如研究家庭状况对青少年犯罪的影响。在某地区未犯罪的青少年为 54000 人，犯罪青少年有 900 人。从全部青少年人口中按 1% 抽样，未犯罪青少年将抽取 540 人，而犯罪青少年只能抽取 9 人。样本量较小，不易提供准确的调查 结果，因而需要扩大犯罪青少年的样本数目，若从 2 人中抽取 1 名，则将抽得 450 人。抽样的结果，因变量中青少年犯罪人数与未犯罪青少年人数在样本中的分布为 1：1.2，即 450 名犯罪青少年与 540 名未犯罪青少年之比，而总体的比例为 1：60，即该地区犯罪青少年人数与未犯罪青少年人数比为 900：54000。为了扩大犯罪 青少年的样本数目，因变量在样本中的分布已不能代表总体中的分布。这种情况下，应按因变量的方向计算百分数。表 9—4 是假定的次数分布，表 9—5 是按自变量即家庭状况方向计算的百分数，表 9—6 是按因变量方向计算的百分数。从表 9—5 看，家庭状况对青少年行为的影响有不尽合理之处。在和睦家庭的 774 名青少年中有 272 名犯罪的，占到 35.14%，这个比率相当高，显然是被夸大的。这是由于在抽样时，扩大了犯罪青少年样本的数目所造成的。按因变量方向计算百分数得到的表 9—6 说明，犯罪青少年来自离异家庭的比率 39.56% 大大高于离异家庭中未犯罪青少年的比率 7.04%。这一结果证明，家庭的状况确实对青少年的行为有影响，家庭的破裂容易引起青少年的犯罪。

表 9—4　　　家庭状况与青少年犯罪

青少年行为	家庭状况		合　计
	离异家庭	和睦家庭	
犯　罪	178	272	450
未犯罪	38	502	540
合　计	216	774	990

表 9—5　　家庭对青少年犯罪行为的影响

青少年行为	家庭状况	
	离　异	和　睦
	%	%
犯　罪	82.41	35.14
未犯罪	17.59	64.86
（总数）	(216)	(774)

表 9—6　家庭对青少年犯罪行为的影响

家庭状况	青少年行为	
	犯　罪	未犯罪
	%	%
离　异	39.56	7.04
和　睦	60.44	92.96
（总数）	(450)	(540)

条件百分表比条件次数表能够提供更多的信息，因此较为有用。但当 $r \times c$ 很大时，百分数的个数会很多，不容易分析两个变量之间的关系。因此，在列联表的相关测量中有许多更实用的方法。

三、PRE 测量法

当研究 X 与 Y 之间的关系，目的是为了用 X 去预测 Y 时，往往希望这一预测的效果是很好的，也就是说是准确的。但事实上难免会出现误差。一般来说，两个变量之间的关系愈强，以一个变量预测另一个变量的误差愈小，也就是减少的预测误差愈多。换句话说，消减的误差有多少，也可以反映变量之间相关程度的强弱。因此，在相关测量中，有必要引进消减误差比例的概念。

相关测量法中有许多种方法，凡是其统计值具有消减误差比

例(Proportionate Reduction in Error)的意义，均称为PRE测量法。

在表9—1中，如果X未知，即不知道妇女的教育水平，仅根据志愿Y所提供的信息预测，那么在390人中预测志愿是“幸福家庭”的最大误差或说全部误差为170人，记作E_1。若X已知，也就是X的分布已知时，被预测错误的人数将为：教育水平高中实际志愿是“理想工作”而被预测为“幸福家庭”的65人；教育水平低中实际志愿是“理想工作”被预测为“幸福家庭”的95人。二者之和是X已知时预测的误差，记作E_2。E_1是全部误差，E_2是利用X与Y之间的联系进行预测时产生的误差，则$E_1 - E_2$就是以X值预测Y值时减少的误差。这个误差与原来全部误差之比，就是消减误差比例，记作PRE。其定义公式为

$$\mathrm{PRE} = \frac{E_1 - E_2}{E_1}$$

如表9—1的资料

$$\mathrm{PRE} = \frac{E_1 - E_2}{E_1} = \frac{170 - (65 + 95)}{170} = 0.0588$$

这一结果表明，利用教育水平去预测妇女的志愿比仅用妇女志愿的边际分布预测可以减少5.88%的误差。简单地说，就是利用教育水平资料预测妇女志愿可减少5.88%的误差。

从PRE的定义公式可以看出，PRE值在0至1之间。PRE的数值愈大，表明X预测Y时能够消减的误差愈多，也就是以X预测Y产生的误差愈小。换句话说，X和Y的相关程度愈高。若$E_2 = 0$，以X预测Y不会产生任何误差，则$\mathrm{PRE} = 1$，X与Y完全相关。若$E_2 = E_1$，以X预测Y的误差等于不以X预测Y的误差，则$\mathrm{PRE} = 0$，X与Y完全无关。表9—1提供的资料，以教育水平预测妇女志愿只能消减5.88%的误差，PRE数值太小，反映X对Y的影响太小，应该寻找其它因素来预测或解释Y。

列联表中的相关测量法中有许多具有消减误差比例的意义，运用时应注意它们对测量层次的要求。

§9.2 列联表中的χ^2检验及相关测量

χ^2 检验可以应用于列联表中资料是否相关的检验。一般适用于两个定类变量之间是否相关的检验，但对定序变量也可以用。对于交互分类表给出的定类，定序变量之间是否独立的检验，χ^2 检验是最有效的。

一、基本方法

对于列联表中的两个变量 X、Y，检验其是否独立，建立的假设组为

H_0： X 与 Y 无关（独立）

H_1： X 与 Y 不独立

运用 χ^2 检验作出判定，需要得到与列联表中实际次数相对应的理论次数。若 X 变量有 $x_1, x_2, \cdots, x_c$ 个值，Y 变量有 $y_1, y_2, \cdots, y_r$ 个数，那么调查获取的条件次数可以排成一个 $r \times c$ 列联表。如表9—2。相对于每一个条件次数 $f_{ij}(i = 1,2,\cdots,r, j = 1,2,\cdots,c)$ 的理论次数即期望次数记作 e_{ij}。则

$$e_{ij} = n\left(\frac{f_{i\cdot}}{n}\right)\left(\frac{f_{\cdot j}}{n}\right) = \frac{f_{i\cdot}f_{\cdot j}}{n} \tag{9.1}$$

式中，$i = 1,2,\cdots,r, j = 1,2,\cdots,c$。条件次数之和与理论预期次数之和相等，都等于总次数，即有

$$\sum_{i=1}^{r}\sum_{j=1}^{c} e_{ij} = \sum_{i=1}^{r}\sum_{j=1}^{c} f_{ij} = n \tag{9.2}$$

若 H_0 成立，则条件次数应是理论的预期次数，也就是说实际次数 f_{ij} 应与理论预期次数 e_{ij} 相等，其差值为0。但测量结果，实际次数 f_{ij} 与理论预期次数 e_{ij} 有差异，这时，可以用其差值的大小来度量两个变量相关的程度。相差愈大，表明 H_0 为真的可能性愈小，即 X 与 Y 无关的可能性愈小。相反，差值愈小，即二者愈接近，H_0 为真的可能性愈大，X 与 Y 之间相关的可能性愈小。为避免 f_{ij} 与 e_{ij} 差值的正负抵消，可以采用差值的平方和，这就是 χ^2 检验中

的统计量 Q。

$$Q = \sum_{i=1}^{r}\sum_{j=1}^{c}\frac{(f_{ij}-e_{ij})^2}{e_{ij}} \tag{9.3}$$

统计量 Q 近似为自由度 $\mathrm{df}=(r-1)(c-1)$ 的 χ^2 分布，在附表 I 中，可以根据给定的显著性水平 α，自由度 df，查找 H_0 为真时的临界值 χ^2_α。若 $Q \geqslant \chi^2_\alpha$，则拒绝 H_0，表明变量 X 与 Y 之间不独立，存在相关。若 $Q < \chi^2_\alpha$，则不能拒绝 H_0，表明变量 X 与 Y 之间独立，不存在相关。

二、应用

【例 9.1】 居住地区对电风扇型式的需求是否有影响

对城镇三类不同地区：热带、温带，寒带的居民进行家电需求量的抽样调查，结果如表 9—7。分析居住地区是否与电风扇的需求型式有关。

表 9—7 不同地区居民对电风扇型式的需求量 单位：万台

电风扇型式	居住地区			合计
	热带	温带	寒带	
台式	14	30	4	48
落地式	67	105	60	232
台地式	30	13	14	57
合计	111	148	78	337

分析：表 9—7 是一个 $r=3$，$c=3$ 的 3×3 列联表。研究两个变量，居住地区与电风扇需求型式之间是否存在相关，可以采用 χ^2 检验。

根据(9.1)式计算与各个条件次数相对应的理论预期次数 $e_{ij}(i=1,2,3; j=1,2,3)$，如

$$e_{11}=\frac{f_{1\cdot}f_{\cdot 1}}{n}=\frac{48(111)}{337}=15.81$$

$$e_{12}=\frac{f_{1\cdot}f_{\cdot 2}}{n}=\frac{48(148)}{337}=21.08$$

类似地依次计算理论预期次数，结果如表 9—8。

表 9—8　不同地区对电风扇型式需求量的理论次数　单位：万台

电风扇型式	居住地区			合计
	热带	温带	寒带	
台式	15.81	21.08	11.11	48.00
落地式	76.42	101.88	3.70	232.00
台地式	18.77	25.04	13.19	57.00
合计	111.00	148.00	78.00	337.00

将表 9—7 和表 9—8 的数据代入(9.3)式计算得到

$$Q=\frac{(14-15.81)^2}{15.81}+\frac{(67-76.42)^2}{76.42}+\frac{(30-18.77)^2}{18.77}+\cdots+\frac{(4-11.11)^2}{11.11}+\frac{(60-53.70)^2}{53.70}+\frac{(14-13.19)^2}{13.19}$$

$$=0.2072+1.1612+6.7189+3.7745+0.0949+5.7819+4.5501+0.7391+0.0497$$

$$=23.0775$$

从上述计算可以看出，按(9.3)式求出统计量 Q 较为麻烦，特别是当 r、c 都较大时。为简化计算，可以采用(9.4)式，它与(9.3)式是一样的，但不用计算理论次数，而是直接利用列联表的实际条件次数和边缘次数计算。公式如下。

$$Q=n\left[\sum_{i=1}^{r}\sum_{j=1}^{c}\frac{f_{ij}^2}{f_{i\cdot}f_{\cdot j}}-1\right] \tag{9.4}$$

根据表 9—7 的数据计算得到

$$Q=337\left[\frac{14^2}{111(48)}+\frac{67^2}{111(232)}+\frac{30^2}{111(57)}+\cdots+\frac{4^2}{78(48)}+\frac{60^2}{78(232)}+\frac{14^2}{78(57)}-1\right]$$

$$=337[1.0684-1]=23.0508$$

按(9.3)式和(9.4)式计算的结果略有差异，这是计算过程中舍取小数造成。

根据自由度 $df = (r-1)(c-1) = (3-1)(3-1) = 4$，显著性水平 $\alpha = 0.05$，在附表 I 中查找，$\chi^2_\alpha = 9.49$。因为 $Q = 23.0508 > \chi^2_\alpha = 9.49$，表明在5%的显著性水平上，数据拒绝 H_0，即居住地区与电风扇的需求型式间不独立，而是存在相关。换句话说，不同居住地区的居民对电风扇的型式有不同的要求。

三、基于 χ^2 值的相关测量法

χ^2 检验利用统计量 Q，可以检验列联表中变量间是否存在相关，但无法测量其相关的程度。在许多的研究中，常称统计量 Q 为 χ^2 值，因此，利用 Q 值计算相关系数，以度量变量间相关程度的方法称之为基于 χ^2 值的相关测量法。列联表中，利用 Q 计算的相关系数主要有以下几种。

1. φ 相关系数

φ 相关系数即 *Phi* 系数。它是由单位频数的 Q 值构成。其计算公式为

$$\varphi = \sqrt{Q/h} \qquad (9.5)$$

式中，Q 是用(9.3)式或(9.4)式计算得到的统计量 Q，n 为列联表的总频数即总次数。

表 9—9　　2 × 2 列联表

Y	X		Σ
	x_1	x_2	
y_1	a	b	$a+b$
y_2	c	d	$c+d$
Σ	$a+c$	$b+d$	$n=a+b+c+d$

φ 值对于 2 × 2 列联表，可以控制在[－1，＋1]这一区间。表

9—9 是一张 2×2 列联表。表中 a、b、c、d 均为条件次数。当变量 X、Y 之间不存在相关时，次数间应有下面的关系，即

$$\frac{a}{a+c}=\frac{b}{b+d}$$

也就是 $ad=bc$

因此，差值 $cd-bc$ 的大小，可以反映变量间关系的强弱。φ 值是利用这一差值来度量变量间相关程度的系数。

根据(9.1)式，与表 9—9 中各实际次数相对应的理论次数 e_{ij} 为

$$e_{11}=\frac{(a+b)(a+c)}{n} \qquad e_{21}=\frac{(a+c)(c+d)}{n}$$

$$e_{12}=\frac{(a+b)(b+d)}{n} \qquad e_{22}=\frac{(b+d)(c+d)}{n}$$

由(9.3)式可得

$$Q=\frac{(a-e_{11})^2}{e_{11}}+\frac{(b-e_{12})^2}{e_{12}}+\frac{(c-e_{21})^2}{e_{21}}+\frac{(d-e_{22})^2}{e_{22}}$$

$$=\frac{n(ad-bc)^2}{(a+b)(c+d)(a+c)(b+d)} \qquad (9.6)$$

代入(9.5)式得到

$$\varphi=\frac{ad-bc}{\sqrt{(a+b)(c+d)(a+c)(b+d)}} \qquad (9.7)$$

当 $ad=bc$ 即 $ad-bc=0$ 时，$\varphi=0$，这时变量 X、Y 之间独立，即不存在相关。若 2×2 表为表 9—10 的形式，$bc=0$，则由(9.7)式计算的 φ 值为 $+1$。这是 X、Y 完全相关的一种情况。若 2×2 列联表如表 9—11 的形式，$ad=0$，则由(9.7)式计算的 φ 值为 -1。这是 X、Y 完全相关的另一种情况。由于列联表中变量的排列是任意的，因此 φ 的符号没有实际意义，即 $\varphi=+1$ 与 $\varphi=-1$ 并不反映变量 X、Y 间相关有什么实质性的差别。通常 $|\varphi|=1$ 时，称作完全相关，也常用 φ^2 进行相关程度的测量。变量 X、Y 完全相关时，列联表的一组对角线上的值必都为零。例如，男、女消费者对某商品的质量评价，一类全部认为好，剩余另一类则全部认为差。一般来说，$|\varphi|$ 小于 1。

表 9—10　　2 × 2 列联表的一种形式

Y	X	
	x_1	x_2
y_1	a	0
y_2	0	d

表 9—11　　2 × 2 列联表的一种形式

Y	X	
	x_1	x_2
y_1	0	b
y_2	c	0

当 $r \times c$ 列联表的 r、c 大于 2 时，φ 值将随 r、c 的增大而增加，φ 值没有上限。这时，不宜反映相关程度的高低。

2. 列联相关系数

列联相关系数是 Pearson 首创的，因此也称作皮尔逊(Pearson)列联系数或 C 系数。它是对 φ 系数的改进，计算公式为

$$C = \sqrt{\frac{Q}{Q + n}} \tag{9.8}$$

当列联表中两个变量相互独立时，$C = 0$。若两个变量存在相关，则 C 值随 r、c 的大小而变化。但它永远小于 1。

列联系数的显著性检验。利用样本数据计算的相关系数，表明在该样本中两个变量间相关的程度，但是否意味着该样本所代表的总体中也存在着这样的相关，需要进行显著性检验。对列联系数 C 的显著性检验，不是利用 C 的抽样分布，而是利用统计量 Q。因为在计算 C 值的过程中先计算了 Q 值，而且 Q 值可以作为 C 值显著性的一个简单的和适合的标志。也就是说，只要检验了 Q 值的显著性，就等于检验了由该 Q 值计算的 C 值的显著性。检验方法同本节第一部分，不再赘述。若样本的 Q 值是显著的，表明总体中两个变量之间存在着相关，列联系数 C 能够表示相关的程度，或说，以

列联系数C值表示总体的相关程度是合适的。

列联系数的局限。C系数计算简便，广泛被使用，但它有几个局限。第一，变量完全相关时，C值不等于1。一般来说，一个令人满意的相关系数至少应满足两个特点：变量完全不相关时，系数等于0；变量完全相关时，系数等于1。列联系数只具备第一个特点，而不具备第二个。第二，两个列联系数不能比较，除非是从同样大小的列联表获得的数据。C值的大小与r、c的大小有关。当$r=c$时，若两个变量完全相关，$C=\sqrt{(r-1)/r}$①。对于2×2表，变量完全相关时，$C=\sqrt{(2-1)/2}=0.7071$；对于3×3表，则有$C=\sqrt{(3-1)/3}=0.8165$。C值随着r、c的增大而变大，因此，由不同r、c的列联表计算的列联系数不能比较。第三，为保证C系数的正确应用，数据必须符合χ^2检验计算的要求，也就是理论预期次数不能小于5。若有的组理论预期次数过小，应与相邻组合并。这样做会使分组变得粗糙，所以抽取样本时应适当增加数据的个数，以避免将组合并。第四，列联系数不能直接与其它相关系数比较，例如Spearman R、Kendall T等。

虽然列联系数有这些局限，但由于它适用于最低一个层次的测量尺度，并且对总体的分布形式不作假设，因此具有广泛的适应性。

3. 克拉默的V相关系数

克拉默的V相关系数(Cramer's V Coefficient of Association)避免了φ值无上限，C值上限不到1的不足，是一个较为适用的以χ^2值为基础的相关系数。其计算公式为

$$V=\sqrt{\frac{Q}{n\min[(r-1),(c-1)]}} \qquad (9.9)$$

根据(9.5)式$\varphi=\sqrt{Q/n}$可以将(9.9)式写成

①$r=c$时，若两个变量完全相关，由(9.4)式可得$Q=n[r-1]$，代入(9.6)式得到$C=\sqrt{n(r-1)/n(r-1)+n}=\sqrt{(r-1)/r}$。

$$V = \sqrt{\varphi^2/\min[(r-1),(c-1)]} \qquad (9.10)$$

分母中 $\min[(r-1),(c-1)]$ 表示取 $(r-1)$ 和 $(c-1)$ 中的较小者。当两个变量不存在相关时，$V=0$，两个变量完全相关，则 $V=1$。

【例 9.2】 以例 9.1 的数据计算 φ、C、V 值。

解：由于在例 9.1 中，两个变量相互独立的零假设被拒绝，表明变量存在相关。利用统计量 Q 分别计算 φ、C、V 值，以测量相关的程度。$Q=23.0508$，$n=337$，$\min[(r-1),(C-1)]=2$，可得到

$$\varphi = \sqrt{Q/n} = \sqrt{23.0580/337} = 0.2615$$

$$C = \sqrt{Q/Q+n} = \sqrt{23.0580/23.0580+337} = 0.2250$$

$$V = \sqrt{Q/n\min[(r-1),(c-1)]}$$

$$= \sqrt{23.0508/337(2)} = 0.1849$$

这三个系数的值愈大，表明两个变量的相关程度愈高。在 $r>2$，$c>2$ 时，φ 值无上限，可能大于 1，因而 $\varphi=0.2615$ 不能认为较大；C 值小于 φ 值，而 C 值的大小与列联表的大小有关，如前所述，对 3×3 表，当两个变量完全相关时，也就是 C 的上限值为 0.8165，相对于 0.8165，$C=0.2530$ 不算太小；V 值在 0 与 1 之间，是一个较好的测量相关程度的系数，对于 $V=1$ 来说，$V=0.1849$ 比较小。因而居住地区与电风扇需求型式的相关程度不够高。这三种相关系数 φ、C、V 都没有消减误差比例的意义，且都仅适用于定类变量，所以其应用受到很大限制。

§9.3 列联表的 PRE 测量法

在实际研究中，仅仅研究变量间相关程度是不够的，往往需要利用变量间的相关关系，从一个变量去预测另一变量。也就是说，在测量相关时，能够得知进行预测将消减多大比例的误差。因此，PRE 测量法比基于 χ^2 值的测量法更有意义。列联表的 PRE 测量法不受测量层次的限制，这也使它比基于 χ^2 值的测量应用更广

泛。列联表的 PRE 测量法主要有以下几种。

一、Lambda 相关测量法

Lambda 相关测量法是以 λ 系数测定变量间相关程度的方法，适用于定类变量间的测量。当一个定序变量与一个定类变量间测定相关程度时，也可以使用 Lambda 测量法，这时，是将定序变量视为定类变量。虽然这样会损失定序变量所提供的等级之分的信息，但却使统计分析易于进行。

1. 非对称形式的 Lambda 相关测量

当研究的两个变量间存在某种因果关系，自变量 X 影响因变量 Y 的变化，而 Y 不会影响 X，这种情况称为非对称关系。例如，研究出生时间和智商的关系，一般认为出生时间对智商会有影响，而智商不会影响出生时间。非对称形式的 Lambda 测量的相关系数以 λ_{yx} 表示，下标 yx 表明 X 是自变量，Y 是因变量，λ 值度量 X 对 Y 的影响程度，具有消减误差比例的意义。

根据

$$\mathrm{PRE} = \frac{E_1 - E_2}{E_1} \tag{9.10}$$

和 $r \times c$ 列联表表 9—2，可以得到 λ_{yx} 的计算公式。

E_1 的定义。当 Y 与 X 的关系未知时，欲预测 Y 的值，唯一的依据是 Y 的边缘分布，即表 9—2 中的列边缘次数 $f_{i\cdot}(i = 1, 2, \cdots, r)$。定类变量的代表值是反映集中趋势的众数，因此，预测 Y 值时的最好办法是以列边缘次数的众数作为 Y 每一个值的预测值。若以 M_y 表示 Y 变量的众数，则以众数预测 Y 值的误差 E_1 为

$$E_1 = n - M_y \tag{9.11}$$

式中，n 为总次数。

E_2 的定义。当 Y 与 X 的关系已知时，预测 Y 的每一个值，则与 X 的值有关。这时，以在某一 X 值的条件下，Y 的众数去预测该条件的 Y 值是误差最小的。以 m_y 表示在每个 X 值条件下 Y 的众数，

则以条件次数中的众数进行预测的误差 E_2 为

$$E_2 = n - \Sigma m_y \tag{9.12}$$

式中，Σm_y 表示各列中 Y 的众数之和。按表 9—2 计算，求和号应为 $\sum_{j=1}^{C}$。将(9.11)式和(9.12)式代入(9.10)式，可以得到非对称形式的的 λ 系数

$$\lambda_{yx} = \frac{\Sigma m_y - M_y}{n - M_y} \tag{9.13}$$

【例 9.3】 出生季节与智商相关程度的分析

出生时间对智力发育不够正常有多大的影响，为此进行了调查。对两组智力不够正常的人的调查结果如表 9—12。

表 9—12 **出生季节与智商**

智商(Y)	出生季节(X)				合计
	春季	夏季	秋季	冬季	
较低	18	29	18	12	77
很低	20	13	16	20	69
合计	38	42	34	32	146

分析：这是两个定类变量间相关程度的分析。因为考察的是出生时间对智力发育是否有影响，所以是非对称关系，可以运用 λ_{yx} 系数度量相关程度。

由表 9—12 可知，$M_y = 77$，春季出生人的智商的众数为 20，夏季出生人的智商的众数为 29，类似地得到秋季 Y 的众数是 18，冬季 Y 的众数是 20，于是 $\Sigma m_y = 20 + 29 + 18 + 20 = 87$，$n = 146$。将上述值代入(9.13)式得到

$$\lambda_{yx} = \frac{87 - 77}{146 - 77} = 0.1449$$

这一结果表明，用出生季节去解释或预测智商，可以减少 14.49% 的误差。

λ_{yx} 表示 Y 受 X 影响的程度，也就是以 X 变量解释或预测 Y 变量时，减少的误差。有时，也可以计算 λ_{xy}。它表示 X 受 Y 影响的程

度。

2. 对称形式的 Lambda 测量

若研究的两个变量 X、Y 之间互相影响，或研究的目的不在于考察哪个变量是自变量，哪个是因变量，则这种情况称为对称关系。例如从事体育活动和收听体育节目之间往往是相互影响的关系。对称形式的 Lambda 测量是计算 λ 系数。

对应于(9.13) 式，可以有

$$\lambda_{xy} = \frac{\Sigma m_x - M_x}{n - M_x} \qquad (9.14)$$

当不考察变量 X 与 Y 哪个为自变量，哪个为因变量时，测度二者之间的相关，可以同时计算 λ_{xy} 和 λ_{xy}，取其平均值作为相关系数，即有

$$\lambda = \frac{n - M_y}{(n - M_y) + (n - M_x)} \cdot \lambda_{yx} + \frac{n - M_x}{(n - M_y) + (n - M_x)} \cdot \lambda_{xy}$$

$$= \frac{\Sigma m_x + \Sigma m_y - (M_x + M_y)}{2n - (M_x + M_y)} \qquad (9.15)$$

λ 是 λ_{yx} 与 λ_{xy} 加权平均的结果。

【例 9.4】 性别和教育程度的相关分析

在某市随机抽取 1200 人调查，结果如表 9—13。

表 9—13　　　　性别和教育程度

教育程度	性别		合计
	男	女	
小学及以下	40	42	82
初　中	182	176	358
高中或中专	204	208	412
大专及以上	200	148	348
合　计	626	574	1200

分析：性别和教育程度可视为相互有关的两个变量。教育程度是定序变量，这里将其视为定类变量，即只考虑分为 4 类，而不考虑其等级。研究两个定类变量间对称关系的相关程度，可以运用 λ 系

的条件次数代入(9.4)式得到

$$Q = 146(0.1107 + 0.1526 + 0.2600 + 0.0583 + 0.1238 + 0.1091 + 0.0584 + 0.1812 - 1)$$
$$= 146(1.0541 - 1)$$
$$= 7.8986$$

根据给定的显著性水平$\alpha = 0.05$，自由度$df = (r-1)(c-1) = 3$，查附表Ⅰ，得到临界值$\chi^2_\alpha = 7.82$。由于$Q = 7.8986 > \chi^2_\alpha = 7.82$，表明数据在5%的显著性水平上，拒绝$H_0$，即在这些智力不够正常的人群中，出生季节与智商存在着相关。但从Q与χ^2_α的数值看，这一结论较为勉强。

从表9—12提供的数据看，夏季和冬季出生的人智商差异较大，将2×4列联表变更为2×2表，仅就夏，冬季的资料分析。表9—15是出生季节与智商的2×2表。

表9—15　出生季节与智商

智商	出生季节		合计
	夏季	冬季	
较低	29	12	41
很低	13	20	33
合计	42	32	74

分析：为研究出生季节与智商的相关程度，应先检验在这群智商不够正常的人中，出生季节与智商是否存在相关。若存在相关，再计算相关系数，度量相关的程度。

H_0：　出生季节与智商无关

H_1：　出生季节与智商相关

将表9—15的数据代入(9.4)式得到

$$Q = 74(0.4884 + 0.1219 + 0.1098 + 0.3788 - 1)$$
$$= 74(1.0989 - 1)$$
$$= 7.3186$$

若以显著性水平 $\alpha = 0.05$，$df = (r-1)(c-1) = 1$ 查附表 I，得到 $\chi^2_\alpha = 3.84$。显然，$Q = 7.3186 > \chi^2_\alpha = 3.84$，数据在5%的显著性水平上拒绝 H_0。当 $\alpha = 0.01$，$df = 1$ 时，$\chi^2_\alpha = 6.64$，数据在1% 的显著性水平上也拒绝 H_0，表明出生季节与智商之间存在着相关。这一结论有足够的说服力。

利用表 9—15 的数据计算 λ_{yx} 为

$$\lambda_{yx} = \frac{(29 + 20) - 41}{74 - 41}$$
$$= 0.2424$$

$\lambda_{yx} = 0.2424$ 表明用出生季节去解释智商的不同，可以减少24.24% 的预测误差。这显然高于例 9.3 的结果。结合条件百分表可以得知，在这些智力不够正常的人当中，夏季出生的人智商高于冬季出生的人，也就是说夏、冬季对智商的影响较大。

【例 9.6】 利用表 9—13 的数据检验某市的性别与教育程度是否相关。

分析：将教育程度视为定类变量，可以运用 χ^2 检验。建立假设组为

H_0： 性别与教育程度无关

H_1： 性别与教育程度相关

利用表 9—13 的数据，根据(9.4)式计算得到

$$Q = 1200(0.0312 + 0.1478 + 0.1614 + 0.1836 + 0.0375 + 0.1507 + 0.1829 + 0.1097 - 1)$$
$$= 1200(1.0048 - 1)$$
$$= 5.76$$

若 $\alpha = 0.05$，$df = (r-1)(c-1) = 3$，查附表 I，得到 $\chi^2_\alpha = 7.82$，若 $\alpha = 0.10$，$df = 3$，查附表 I，得到 $\chi^2_\alpha = 6.25$。比较 Q 及两个临界值可知，数据在 5% 的显著性水平上，甚至在 10% 的显著性水平上都不能拒绝 H_0，因此，该市的性别与教育程度无关。这一检验表明，例 9.4 中计算的 $\lambda_{yx} = 0$、$\lambda = 0.004$ 确实反映了性别与教育

程度的相关情况。

一般来说，假设检验关心的是总体是否存在相关，也就是以样本的资料来推断变量在总体中是否相关，而相关测量法是研究的变量在样本中的相关程度和方向。假设检验要求样本必须是随机抽样获得，而相关测量法则既可用于随机抽样的样本，也可用于非随机抽样的样本。当样本是随机抽样获得时，一般应先进行假设检验，以判定变量在总体中是否存在相关。若存在相关，再测度相关的程度。

二、Goodman-Kruskal Tau 相关测量法

Goodman-Kruskal Tau 相关测量法是由古德曼和克鲁斯卡尔创造的，采用 Tau 系数测定两个定类变量间的相关程度。Tau 系数是对 Lamhda 系数的改进。它不再用众数对 Y 进行预测，而是利用边缘次数提供的比例进行预测。

1. 非对称形式的 Tau 相关测量

若两个定类变量 X、Y 之间存在因果关系，自变量为 X，因变量为 Y，测量其相关程度，可以采用 Tau 系数，也记作 τ 系数。由于是非对称关系的相关测量，因此，也记作 τ_y。Tau 相关测量法具有消减误差比例的意义。

E_1 的定义。当 X、Y 的关系未知时，可以用表 9—2 中的边缘次数比例预测 Y 值。边缘次数比例 $f_{1\cdot}/n$ 是 Y_1 值的概率，若以 $f_{1\cdot}/n$ 预测 Y_1 值，那么误差将是 $f_{1\cdot}(1-f_{1\cdot}/n)$。因为 $f_{1\cdot}/n$ 是 Y_1 出现的概率，而 $1-f_{1\cdot}/n$ 就是 Y_1 值不出现的概率。同样的道理，预测 Y_2 值的误差将是 $f_{2\cdot}(1-f_{2\cdot}/n)$，预测 Y_3 值的误差为 $f_{3\cdot}(1-f_{3\cdot}/n)$ 等等。这样，X、Y 关系未知时，以边缘次数比例预测 Y 值的全部误差 E_1 为

$$
\begin{aligned}
E_1 &= f_{1\cdot}(1-f_{1\cdot}/n)+f_{2\cdot}(1-f_{2\cdot}/n) \\
&\quad +\cdots f_{r\cdot}(1-f_{r\cdot}/n) \\
&= f_{1\cdot}+f_{2\cdot}+\cdots+f_{r\cdot}-[(f_{1\cdot}^2+f_{2\cdot}^2+\cdots+f_{r\cdot}^2)/n]
\end{aligned}
$$

$$= n - \sum_{i=1}^{r} f_{i\cdot}^2 / n \tag{9.16}$$

E_2 的定义。当 X、Y 关系已知时，根据表 9—2 中的条件次数比例预测 Y 值。条件次数比例 $f_{ij}/f_{\cdot j}$ 是 Y_i 值出现的概率，而 $1 - f_{ij}/f_{\cdot j}$ 是 Y_i 值不出现的概率，所以用条件次数比例 $f_{ij}/f_{\cdot j}$ 预测 Y_i 值的误差为 $f_{ij}(1 - f_{ij}/f_{\cdot j})$。这样，当 X、Y 关系已知时，以条件次数比例预测 Y 的总误差 E_2 为

$$\begin{aligned}
E_2 &= f_{11}(1 - f_{11}/f_{\cdot 1}) + f_{12}(1 - f_{12}/f_{\cdot 2}) \\
&\quad + \cdots + f_{rc}(1 - f_{rc}/f_{\cdot c}) \\
&= (f_{11} + f_{21} + \cdots + f_{r1}) - \sum_{i=1}^{r} f_{i1}^2/f_{\cdot 1} \\
&\quad + (f_{12} + f_{22} + \cdots + f_{r2}) - \sum_{i=1}^{r} f_{i2}^2/f_{\cdot 2} \\
&\quad + \cdots + (f_{1c} + f_{2c} + \cdots + f_{rc}) - \sum_{i=1}^{r} f_{ic}^2/f_{\cdot c} \\
&= n - \sum_{j=1}^{c} \sum_{i=1}^{r} f_{ij}^2/f_{\cdot j}
\end{aligned} \tag{9.17}$$

将(9.16)式和(9.17)式代入(9.10)式，得到 Tau 测量系数为

$$\begin{aligned}
\tau_y &= \frac{E_1 - E_2}{E_1} = \frac{n - \sum_{i=1}^{r} f_{i\cdot}^2/n - n + \sum_{j=1}^{c}\sum_{i=1}^{r} f_{ij}^2/f_{\cdot j}}{n - \sum_{i=1}^{r} f_{i\cdot}^2/n} \\
&= \frac{\sum_{j=1}^{c}\sum_{i=1}^{r} f_{ij}^2/f_{\cdot j} - \sum_{i=1}^{r} f_{i\cdot}^2/n}{n - \sum_{i=1}^{r} f_{i\cdot}^2/n}
\end{aligned} \tag{9.18}$$

【例 9.7】 城乡地区与购买电风扇信息途径的相关测量

随机从城镇、乡村两个地区抽取 10800 户家庭调查，结果如表 9—16。

分析：表中提供的数据是两个定类变量的值，测定城乡地区与购买电风扇信息途径的相关程度，是为考察城乡不同地区信息途径是否不同，对两个不同地区是否应采用不同的宣传形式。所以，自变量是不同地区，因变量是不同的信息途径。这是非对称的关系，可以采用 Tau 相关测量法。运用(9.18)式计算得到

表 9—16　　城乡的购买电风扇信息途径

信息途径	地区		合计($f_{\cdot j}$)
	城镇	乡村	
熟人推荐	3191	1120	4311
商店推荐	2370	1183	3553
电视广告	1286	528	1814
其他媒介	935	187	1122
合计($f_{\cdot j}$)	7782	3018	10800

$$\tau_y=\frac{\left(\frac{3191^2}{7782}+\frac{2370^2}{7782}+\cdots+\frac{187^2}{3018}\right)-\frac{(4311^2+3553^2+\cdots+1122^2)}{10800}}{10800-(4311^2+3553^2+1814^2+1122^2)/10800}$$

$$=\frac{3338.416-3310.9269}{10800-3310.9269}$$

$$=0.0037$$

$\tau_y=0.0037$ 表明用城乡地区去解释购买电风扇的信息途径，只能消减 0.37% 的预测。因此，试图用城乡不同地区对促销采用的宣传形式作出决策是不大合适的。

2. Tau 相关测量法的特点。

τ_y 的取值范围。当 X、Y 之间完全无关时，$\tau_y=0$；当 X、Y 完全相关时，$\tau_y=1$。一般来说，τ_y 值越接近于 1，表明 X、Y 间的相关程度越高，以 X 去解释 Y，能够消减的预测误差比例越大。

τ_y 值的非对称性。τ_y 值表明用 X 去解释 Y 时能够消减的误差比例。若以 Y 解释 X，即 Y 为自变量，X 为因变量，则应计算 τ_x。一般 $\tau_x\neq\tau_y$ 所以是非对称的。但对于 2×2 表，$\tau_x=\tau_y$。

【例 9.8】　利用表 9—16 计算 τ_x

分析：τ_y 表明不同地区对购买电风扇信息途径的影响，τ_x 表明购买电风扇信息途径不同是否与不同地区有关。根据(9.18) 式可以得到

$$\tau_x=\frac{\left(\frac{3191^2}{4311}+\frac{1120^2}{4311}+\cdots+\frac{187^2}{1122}\right)-\frac{7782^2+3018^2}{10800}}{10800-(7782^2+3018^2)/10800}$$

$$= \frac{52.7081}{4349.2733}$$
$$= 0.0121$$

$\tau_x = 0.0121$ 大于 $\tau_y = 0.0037$，表明以 X 解释 Y 和以 Y 解释 X，能够消减的误差比例不同，这表明 X 变量中所包含的关于 Y 的信息与 Y 变量中包含的关于 X 的信息并不是等量的。

τ_y 与 λ_{yx} 的比较。τ_y 与 λ_{yx} 都测定非对称关系的定类变量间相关程度。但其依据不同。λ_{yx} 以众数作为预测依据，τ_y 则利用了列联表中的每一个条件次数。相对来说，τ_y 比 λ_{yx} 更精细。但 λ_{yx} 比 τ_y 计算简便。一般情况，当列联表中众数频次比较突出时，宜采用 λ_{yx} 系数；变量分类较多，众数频次与非众数频次差异不悬殊时，宜采用 τ_y 系数。若列联表中，众数频次集中在同一横行，λ_{yx} 等于零，必须运用 τ_y 系数来反映变量间的相关程度。

【例 9.9】 利用表 9—13 计算 τ_y

分析：在例 9.4 中，计算 $\lambda_{yx} = 0$，这是因为在不同性别下的教育程度众数频次均在"高中或中专"这一行。这种情况下，用 λ_{yx} 测度相关程度不大合适，可以采用 τ_y 系数。利用表 9—13 的数据，根据(9.18)式，得到

$$\tau_y = \frac{\left(\frac{40^2}{626} + \frac{182^2}{626} + \cdots + \frac{148^2}{570}\right) - \frac{82^2 + 358^2 + 412^2 + 348^2}{1200}}{1200 - (82^2 + 358^2 + 412^2 + 348^2)/1200}$$
$$= \frac{1.6381}{845.22}$$
$$= 0.0019$$

$\tau_y = 0.0019$，表明用性别去解释教育程度，可以消减 0.19% 的预测误差。这一结果与 λ_{yx} 不同。

3. 显著性检验

与 Lambda 相关测量法一样，在运用 Tau 相关测量法测定两个变量间的相关程度之前，应先进行显著性检验，以判定总体中两个变量是否存在相关。若确实存在相关，再计算 τ_y，测定相关程度。对于两个定类变量，检验其在总体中是否相关，宜采用 χ^2 检验。

【例 9.10】 年龄是否对影片的评价有影响

随机抽选了 100 人进行调查，了解老、中、青三种不同年龄的人，对某部影片的评价是否存在显著差异。若存在，测定相关程度。调查结果如表 9—17。

分析：研究不同年龄的人对该部影片的评价是否存在显著差异，实际上是检验变量“年龄”和变量“评价”之间是否相互独立。若相互独立，两个变量间无关；若相互不独立，两个变量间存在相关。这一检验采用 χ^2 检验。

H_0： 年龄与评价无关

H_1： 年龄与评价有关

表 9—17 年龄与对影片的评价

对影片评价	年龄			合计
	老年	中年	青年	
评价很高	14	13	13	40
评价一般	18	2	15	35
评价较低	8	15	2	25
合　　计	40	30	30	100

根据表 9—17 数据，代入(9.4)式得到

$$
\begin{aligned}
Q &= 100[0.1225 + 0.2314 + 0.0640 + 0.1408 + \\
&\quad 0.0038 + 0.3000 + 0.1408 \\
&\quad + 0.2143 + 0.0053 - 1] \\
&= 100[1.2229 - 1] \\
&= 22.29
\end{aligned}
$$

若给定 $\alpha = 0.05$，以 $\mathrm{df} = (r-1)(c-1) = 4$ 查附表 I，得到 $\chi^2_\alpha = 9.49$。显然 $Q = 22.29 > \chi^2_\alpha = 9.49$，表明数据在 5% 的显著性水平上拒绝 H_0，年龄与对影片的评价有关。

这里考察的是年龄对评价有无影响，因此是非对称关系。计算

τ_y 来测定相关程度。利用(9.18)式可得到

$$\tau_y = \frac{\left[\frac{14^2+18^2+8^2}{40}+\frac{13^2+2^2+15^2}{30}+\frac{13^2+15^2+2^2}{30}\right]-\frac{40^2+35^2+25^2}{100}}{100-(40^2+35^2+25^2)/100}$$

$$= \frac{6.6334}{65.5}$$

$$= 0.1013$$

$\tau_y = 0.1013$ 表明年龄与评价之间存在较弱的相关。以年龄去解释对该部影片的评价只能消减 10.13% 的误差。

【例 9.11】 家庭规模与经济状况的相关分析

从某地区随机抽取 100 户家庭调查，结果如表 9—18。

表 9—18　　家庭规模与经济状况

经济状况	家庭规模		合计
	大	小	
高收入	2	44	46
中等收入	32	10	42
低收入	10	2	12
合　计	44	56	100

分析：要进行家庭规模与经济状况的相关分析，首先应利用调查获得的样本数据，判定变量“家庭规模”与变量“经济状况”在总体中是否相关，总体即该地区的全部家庭，因为样本取自该地区的全部居民户之中。

H_0:　家庭规模与经济状况无关

H_1:　家庭规模与经济状况有关

将“经济状况”变量视作定类变量，采用 χ^2 检验判定。根据表 9—18 的数据，运用(9.4)式到

$$Q = 100[3^2/44(46) + 32^2/44(42) + \cdots + 2^2/56(12) - 1]$$

$$= 100[1.5455 - 1]$$

$$= 54.55$$

运用 Gamma 相关测量法测定两个定序变量间的相关程度，其目的是考察根据一个变量的某一等级去预测另一变量的等级时，能消减的误差比例为多大。这需要利用两个变量等级之间的关系。序对是指高、低位次的两两配对。

同序对是 X 变量中的数值与 Y 变量中的数值变化方面一致的序对。表 9—19 是三个职工受教育程度与经济收入的列联表。由表 9—19 看，甲、乙、丙三人在受教育程度和经济收入方面的位次有如下的关系：

受教育程度：　乙 > 甲；乙 > 丙；甲 > 丙

经济收入：　　甲 > 乙；乙 > 丙；甲 > 丙

从上述关系可以看出，乙和丙在受教育程度与经济收入上的位次是一致的，均为乙 > 丙，这是一个同序对。甲和丙也存在这样的一致位次关系，所以也是一个同序对。表 9—19 有两个同序对，记作 $n_s = 2$。

表 9—19　三个职工的受教育程度与经济收入

经济收入	受教育程度		
	高	中	低
高		甲	
中	乙		
低			丙

异序对是两个变量数值变化方向不一致的序对，也就是位次相反的序对。表 9—19 中，甲和乙在受教育程度、经济收入方面的位次不同，受教育程度上乙高，经济收入上却是甲高，这是一个异序对，记作 $n_d = 1$。

同分对是变量的数值等级相同的序对。同分对可以出现在变量 X 中，也可以出现在变量 Y 中，还可以出现在两个变量之间。变量 X 中的同分对记作 T_x，变量 Y 中的同分对记作 T_y，变量 X、Y 之间的同分对记作 T_{xy}。

2. 对称关系的 Gamma 相关测量

Gamma 相关测量法具有消减误差比例的意义，它是利用同序对和异序对来定义系数 G 的。

E_1 的定义。若变量 X、Y 间的关系未知时，以 X 的等级去预测 Y 的等级纯属随机，这时预测正确和错误的概率各为 1/2。若总的序对为 $n_d + n_s$（不计同分对），那么预测 Y 的全部误差 E_1 为

$$E_1 = \frac{1}{2}(n_s + n_d) \tag{9.19}$$

E_2 的定义。当 X、Y 之间的等级关系已知，那么以同序进行预测，即对于序对 (x_i, y_i)、(x_j, y_j)，当 $x_i > x_j$ 时，预测 $y_i > y_j$；当 $x_i < x_j$ 时，预测 $y_i < y_j$。这样预测的结果，可能错误的数目为异序对数目。即有

$$E_2 = n_d \tag{9.20}$$

将(9.19)式和(9.20)式代入(9.10)式可以得到

$$G = \frac{E_1 - E_2}{E_1} = \frac{1/2(n_s + n_d) - n_d}{1/2(n_s + n_d)}$$

$$= \frac{n_s - n_d}{n_s + n_d} \tag{9.21}$$

(9.21)式就是Gamma 相关测量法的系数 G。(9.21)式没有考虑 Y 对 X 的影响，还是 X 对 Y 的影响，因而是对称关系的相关测量。

3. 列联表中 n_s 和 n_d 的计算

列联表中的调查总数目 n 一般都很大，为计算 n_s 和 n_d，应将数据在列联表中按等级顺序排列，如表 9—20 所示。以 f_{11} 为基础分析。f_{11} 是处在最高位次，无论对 X 还是 Y，取 f_{22} 与其配对，则在 X 变量上的等级与 Y 变量上的等级方向一致，构成同序对；同样地，f_{11} 与 f_{33}，f_{11} 与 f_{32}，f_{11} 与 f_{23} 配对，都构成同序对。因此，对 f_{11} 来说，共可得到同序对为

$$f_{11}(f_{22} + f_{33} + f_{23} + f_{32})$$

表 9—20　两个定序变量的 3 × 3 列联表

Y	X		
	高	中	低
高	f_{11}	f_{12}	f_{13}
中	f_{21}	f_{22}	f_{23}
低	f_{31}	f_{32}	f_{33}

从上述分析中可以看出，对 f_{11} 来说，同行、同列的均形成同分对，只有其右下方非同行非同列的才能形成同序对。这一规律对其它位置上的频数都适用。

以 f_{12} 为基础分析，按上面所述的规律，与其同行同列的为同分对，在其左下侧的构不成同序对。因此，对 f_{12} 来说，同序对为

$$f_{12}(f_{23}+f_{33})$$

f_{13} 不可能与其它频数构成同序对。依据上面的规律，f_{21} 的同序对为

$$f_{21}(f_{32}+f_{33})$$

同理，f_{22} 的同序对为 $f_{22}(f_{33})$。第三行不可能与任何一个频数构成同序对，所以同序对的总数为上面 4 类之和，即

$$n_s=f_{11}(f_{22}+f_{33}+f_{23}+f_{32})+f_{12}(f_{23}+f_{33})$$
$$+f_{21}(f_{32}+f_{33})+f_{22}(f_{33})$$

异序对可以采用类似的方法得到，即将同行同列的同分对舍去，再舍去某一频数右下方的同序对，在某一频数左下方的都可构成异序对。表 9—20 中的异序对总数为

$$n_d=f_{13}(f_{21}+f_{31}+f_{22}+f_{32})+f_{12}(f_{21}+f_{31})$$
$$+f_{23}(f_{31}+f_{32})+f_{22}(f_{31})$$

列联表中的同序对、异序对依据上面的准则可以计算得到，将 n_s、n_d 代入(9.21)式即能求出系数 G。

【例 9.12】　经济状况和家庭规模的相关分析

从某地区随机抽取 100 户家庭调查，结果如表 9—21。

分析：表 9—21 是两个定序变量按从高到低的顺序排列的 2 × 3 列联表。研究变量“经济状况”和变量“家庭规模”之间的相关程度，可以采用 Gamma 测量法。

计算同序对

$$n_s = 2(10 + 2) + 32(2) = 88$$

计算异序对

$$n_d = 10(44 + 10) + 32(44) = 1948$$

将 n_s、n_d 代入(9.21) 式得到

$$G = \frac{n_s - n_d}{n_s + n_d} = \frac{88 - 1948}{88 + 1948} = -0.9136$$

表 9—21　　经济状况与家庭规模

家庭规模	经济状况			合计
	高收入	中等收入	低收入	
大	2	32	10	44
小	44	10	12	56

$G = -0.9136$，表明两个变量间存在负相关；$|G| = 0.9136$，接近于 1，表明两个变量之间的相关程度很高。由于 G 系数具有消减误差比例的性质，因此 $|G| = 0.9136$，意味着，以经济状况的相对等级解释家庭规模的相对等级可以消减 91.36% 的误差。因为，G 系数是对称关系的相关测量，因而也可以说，以家庭规模的相对等级解释经济状况的相对等级可以消减 91.36% 的误差。经济状况与家庭规模呈负相关，表明一个变量等级愈高，另一变量等级愈低，即经济收入越高，家庭规模越偏小。

4. Gamma 相关测量法的特点

G 系数的取值范围。当不考虑同分对时，若 $n_d = 0$，即数据都是同序对，则 $G = 1$，若 $n_s = 0$，即数据都是异序对，则 $G = -1$。所以，G 系数的取值在[-1，$+1$] 区间。若数据中以同序对为主，即 $n_s > n_d$，则 $G > 0$，表明两个变量呈正相关；若数据以异序对为主；

$$\varphi = \frac{ad - bc}{\sqrt{(a+b)(c+d)(a+c)(b+d)}}$$
$$= \frac{60(12) - 0}{\sqrt{108(12)(60)(60)}}$$
$$= 0.33$$

计算 Q 系数为

$$Q = \frac{ad - bc}{ad + bc}$$
$$= 1$$

由于在这一研究中，关心的是服用新药能否不患肝炎，而对服用原有药是否能不患肝炎不感兴趣，因此，采用 Q 系数更适宜，也就是说用 Q 系数反映新药与肝炎的关系更合理。

2×2 表中的 Q 系数可以看作 G 系数的特例。

5. Gamma 系数的检验

利用随机样本数据计算的 G 系数，是否能用以推断总体，必须进行统计检验。建立的假设组为

H_0：　总体中 $G = 0$

H_1：　总体中 $G \neq 0$

若需检验 $G > 0$ 或 $G < 0$，也可以建立单侧备择，即

H_0：　$G = 0$　　　H_0：　$G = 0$

H_+：　$G > 0$　　　H_-：　$G < 0$

为判定假设，需要采用随机抽样获得数据，数据至少是定序尺度测量。定义的检验统计量为

$$Z = G\sqrt{\frac{n_s + n_d}{n(1 - G^2)}} \tag{9.23}$$

其中，G 是利用样本数据计算的 G 系数，n_s、n_d 是列联表中的同序对、异序对，n 为样本数据的数目。统计量 Z 是 G 系数标准化的结果，当 $n \geqslant 10$ 时，它近似为标准正态分布。因而可以在附表 N 中，查找到 H_0 成立时，Z 为某数值的概率 P。将 P 与给定的显著性水平 α 比较，若 P 足够小，则数据拒绝 H_0，否则不能拒绝 H_0。利用统

计量 Z 作出判定，也可以在附表 Ⅳ 中，根据给定的显著性水平查找临界值 Z_α，即确定单侧检验的否定域 $|Z| \geq Z_\alpha$，若统计量 Z 的值落入否定域，则拒绝 H_0，否则不能拒绝 H_0。若检验为双侧的，应以 $\alpha/2$ 查找附表 Ⅳ 中相应的临界值 $Z_{\alpha/2}$。

【例 9.14】 某地区的家庭中，家庭规模是否随经济收入升高而变小

利用例 9.12 中某地区家庭状况的调查资料，能否说明该地区的居民家庭，经济收入等级越高，家庭规模越小。

分析：在例 9.12 中，调查的样本是随机抽取的，利用样本数据计算的 $G = -0.9136$，表明家庭规模与经济收入呈现高度的负相关。能否将这一结论推断到该地区居民家庭，应进行检验。由于研究的是经济收入越高，家庭规模越小，因而建立单侧备择，假设组为

$$H_0: \quad G = 0$$
$$H_1: \quad G < 0$$

根据表 9—21 计算 $n_s = 88, n_d = 1948, n = 100, G = -0.9136$，代入(9.23)式，得到

$$Z = (-0.9136)\sqrt{\frac{88 + 1948}{100[1 - (-0.9136)^2]}}$$
$$= -10.1393$$

若要求显著性水平 $\alpha = 0.001$，则在附表 Ⅳ 中可以查到单侧检验的临界值 $Z_\alpha = 3.09$，即否定域为 $|Z| \geq 3.09$。显然统计量 $|Z| = 10.1393$，落入否定域，所以数据拒绝 H_0。若以 $|Z| = 10.1393$ 在附表 Ⅳ 中查找 H_0 成立时的概率可得到 $P \leq 0.0002$。这是一个相对于 $\alpha = 0.001$ 足够小的概率，故数据拒绝 H_0。

研究的结论是：该地区居民家庭的规模与经济收入成反比，家庭规模是随经济收入的升高而变小。

四、Somer's d 相关测量法

Somer's d 相关测量法亦称 d 相关测量法，是通过计算 d 系数

$$= 168 + 110 + 48 + 64$$
$$= 390$$
$$T_y = 38(5 + 4) + 5(4) + 18(16 + 6) + 16(6)$$
$$+ 4(4 + 5) + 4(5)$$
$$= 342 + 20 + 396 + 96 + 36 + 20$$
$$= 910$$

根据(9.24)式可以计算得到

$$d_{yx} = \frac{n_s - n_d}{n_s + n_d + T_y} = \frac{1475 - 390}{1475 + 390 - 910}$$
$$= 0.3910$$

$d_{yx} = 0.3910$ 表明父母的教育态度对子女的学习成绩有一定的影响，二者成正比即父母的教育态度愈好，相对来说，子女的学习成绩要高，以父母教育态度好坏预测说明子女学习成绩的高低可以削减近 40% 的误差。

d_{yx} 系数与 G 系数都应用了同序对 n_s 和异序对 n_d，但 G 系数未考虑同分对，而 d_{yx} 系数则考虑了。一般情况下，研究的两个变量是对称关系，即不研究哪一个变量为自变量，哪一个为因变量，采用 G 系数测定相关程度；若研究的两个变量需要区分自变量、因变量时，应采用 d_{yx} 系数或 d_{xy} 系数测定相关程度。在社会现象的研究中，往往在区分自变量、因变量的情况下，也采用 G 系数，虽这样不够严谨，但也大致可以接受。

3. d 系数的检验

若样本是随机抽样选取的，那么能否用样本资料来推断两个变量在总体中是否相关需要进行统计检验。在 d 系数的计算中，分子是同序对与异序对的差值，即有

$$S = n_s - n_d$$

也就是说，S 与总体是否存在等级相关有关。若 $S = 0$，则 X 与 Y 之间不存在等级相关；若 $S \neq 0$，则存在等级相关。利用 S 因子的数值，可以对总体是否存在等级相关进行检验。

根据研究问题的需要，可以建立双侧备择，也可以建立单侧备择。双侧备择为

H_0：　总体中 $S=0$

H_1：　总体中 $S\neq 0$

单侧备择为

H_0：　$S=0$　　　H_0：　$S=0$

H_{+}：　$S>0$　　　H_{-}：　$S<0$

为对假设作出判定，需要计算检验统计量 Z。Z 的计算公式为

$$Z=\frac{S'}{S_e} \tag{9.26}$$

式中，S' 是 S 的修正值

$$S'=|S|-\frac{n}{2(r-1)(c-1)}$$

这里，r、c 分别是列联表的行数、列数，n 为样本数据的个数。(9.26) 式中的 S_e 是 S' 的标准误差，其计算公式为

$$S_e=\sqrt{\frac{A_2B_2}{n-1}-\frac{A_2B_3+A_3B_2}{n(n-1)}+\frac{A_3B_3}{n(-1)(n-2)}}$$

这里，A_2 是 X 变量边缘次数中，每 2 个频次乘积之和，A_3 是每 3 个频次乘积之和；B_2 是 Y 变量边缘次数中，每 2 个频次乘积之和，B_3 是每 3 个频次乘积之和。

检验统计量 Z 近似正态分布，可以在附表 Ⅳ 中，查得与给定的显著性水平相对应的临界值 Z_α，将 Z 与 Z_α 比较，若 $|Z|\geqslant Z_\alpha$，则拒绝 H_0，否则不能拒绝 H_0。若检验为双侧的，查附表 Ⅳ 时，应以 $\alpha/2$ 查找相应的临界值 $Z_{\alpha/2}$。

【例 9.16】　子女的学习成绩高低是否随父母的教育态度好坏而变化

沿用例 9.15 的调查结果，分析该地区的子女学习成绩与父母教育态度的关系。

分析：表 9—24 的数据是在某地区随机抽样调查得到，以样本数据计算的 $d_{yx}=0.3910$，这一结论是否能用来推断总体，要进行

检验。利用上面介绍的 Z 检验法，建立的假设组为

H_0: 总体中 $S=0$

H_+: 总体中 $S>0$

由表 9—24 可知：$n=100, r=3, c=3$，所以

$$S'=|1475-390|-\frac{100}{2(3-1)(3-1)}$$
$$=1072.5$$

将表 9—24 中 X、Y 的边缘次数计算出来列入表 9—25 中，可以得到

$$A_2=60(25+15)+25(15)$$
$$=2400+375$$
$$=2775$$
$$A_3=60(25)(15)$$
$$=22500$$
$$B_2=47(40)+47(13)+40(13)$$
$$=3011$$
$$B_3=47(40)(13)$$
$$=24440$$

表 9—25　**父母教育态度与子女学习成绩**

子女学习成绩(Y)	父母教育态度(X)			合计
	好	中	差	
高	38	5	4	47
中	18	16	6	40
低	4	4	5	13
合计	60	25	15	100

于是

$$S_e=\sqrt{\frac{2775(3011)}{100-1}-\frac{2775(24440)+22500(3011)}{100(100-1)}+\frac{22500(24440)}{100(100-1)(100-2)}}$$
$$=266.963$$

根据(9.26)式可得

$$Z = \frac{S'}{S_e} = \frac{1072.5}{266.963}$$
$$= 4.0174$$

若给定显著性水平 $\alpha = 0.001$，则在附表 Ⅳ 中可查得临界值为 $Z_\alpha = 3.09$。否定域为 $|Z| \geqslant 3.09$。显然，$Z = 4.0174$ 已落入否定域，数据拒绝 H_0。结论是：该地区子女的学习成绩与父母的教育态度成正比，父母教育态度愈好，子女学习成绩愈高。

五、相关比率测量法

相关比率(correlation ratio)亦称eta平方系数，写作eta^2或简写为 E^2。它用来测定一个定类变量(X)或一个定序变量(X)与一个定距变量(Y)之间的相关程度．有时也记作 η^2。

1.不对称关系的 E^2 系数

将定类变量或定序变量作为 X，定距变量作为 Y，那么 E^2 的计算公式为

$$E^2 = \frac{\Sigma n_i \bar{y}_i^2 - n\bar{y}^2}{\Sigma y_i^2 - n\bar{y}^2} \tag{9.27}$$

式中：　n_i：　自变量的每组观察值数目；

n：　样本观察值总数，$n = \Sigma n_i$；

y_i：　因变量的数值；

$\bar{y}_i$：　第 i 组的因变量均值；

$\bar{y}$：　因变量的均值。

E^2 具有削减误差比例的意义。当 X、Y 的关系未知时，最好的预测量以因变量 Y 的样本均值预测因变量的每一个值，这时的全部误差 E_1 为

$$E_1 = \Sigma(y_i - \bar{y})^2 \text{①} \tag{9.28}$$

若 X、Y 的关系已知，这时以与 x_i 对应的因变量的样本均值 $\bar{y}_i$ 作为

① 采用平方值是为避免离差的正负抵消。

该组的预测值最好，误差将为组内离差平方和

$$E_2 = \Sigma(y_i - \bar{y}_i)^2 \qquad (9.29)$$

根据 PRE 的公式可以得到

$$\begin{aligned} \text{PRE} &= \frac{E_1 - E_2}{E_1} = \frac{\Sigma(y_i - \bar{y})^2 - \Sigma(y_i - \bar{y}_i)^2}{\Sigma(y_i - \bar{y})^2} \\ &= \frac{\Sigma n_i \bar{y}_i^2 - n\bar{y}^2}{\Sigma y_i^2 - n\bar{y}^2} \\ &= E^2 \end{aligned}$$

【例 9.17】 居民的收入与文化消费支出是否相关。

在某地区随机抽选了 20 名居民，对其收入与用于文化娱乐消费的支出进行调查，结果如表 9—26。

表 9—26　　居民收入与文化消费支出

		收入(X)		
		高收入	中等收入	低收入
文化消费支出（元）	(Y)	4.20	3.20	3.10
		3.00	2.90	2.80
		2.90	2.70	2.20
		3.20	2.10	1.10
		2.70	1.90	2.00
		3.10	3.10	2.50
		2.30	2.80	

分析：由于居民收入是按定序尺度测量，文化消费支出是定比尺度测量，测定其相关程度，可以采用 E^2 系数。

由表 9—26 可以计算得到：

$$n_1 = 7, n_2 = 7, n_3 = 6, n = n_1 + n_2 + n_3 = 20$$

$$\bar{y}_1 = 3.0571, \bar{y}_2 = 2.6714, \bar{y}_3 = 2.2833, \bar{y} = 2.69$$

$$\Sigma y^2 = 152.64$$

代入(9.27)式有

相关比率的检验统计量 F 可以不先计算 E^2，而根据方差分析的方法计算。单因素方差分析的基本原理是将因变量的全部离差平方和 $\Sigma(Y-\overline{Y})^2$ 即 E_1 分解为两部分：削减的误差平方和和剩余的误差平方和。削减的误差平方和即 $E_1 - E_2$，也就是由于引进自变量 X 来解释 Y 的变化所能削减的误差平方和，称为组间离差平方和，记作 BSS(between-groups sum of sguares)，由(9.28)式、(9.29)式知：

$$\begin{aligned} BSS &= \Sigma(y-\overline{y})^2 - \Sigma(y-\overline{y}_i)^2 \\ &= \Sigma n_i \overline{y}_i^2 - n\overline{y}^2 \\ &= \Sigma n_i(\overline{y}_i - \overline{y})^2 \end{aligned}$$

剩余误差平方和即 E_2，是用自变量 X 无法解释 Y 的那部分离差平方和，称为组内离差平方和，记作 WSS(Within-groups sum of sguares)，由(9.29)式知：

$$WSS = \Sigma(y-\overline{y}_i)^2$$

于是就可以构造检验统计量 F。BSS 和 WSS 分别除以各自的自由度 $k-1$ 和 $n-k$，得到被削减的方差和剩余的方差，两个方差之比，就是 F 值，即有

$$F = \frac{BSS/(k-1)}{WSS/(n-k)} = \frac{\Sigma n_i(\overline{y}_i - \overline{y})^2/(k-1)}{\Sigma(y-\overline{y}_i)^2/(n-k)} \qquad (9.31)$$

(9.30)式与(9.31)式是等价的。仍用表9—26的数据按(9.31)式计算检验统计量 F 为

$$F = \frac{[7(3.0571-2.69)^2 + 7(2.6714-2.69)^2 + 6(2.2833-2.69)^2]/(3-1)}{[(4.20-3.0571)^2 + (3.00-3.0571)^2 + \cdots + (2.50-2.2833)^2]/(20-3)}$$
$$= \frac{1.9381/2}{5.9799/17}$$
$$= 2.7549^{①}$$

在列联表的PRE测量法中，上述几种较为常用。在选择时，应

① 由(9.31)式计算的 F 值与(9.30)式计算的 F 值差异是由于小数位保留引起。

首先考虑变量的测量层次，其次才是关系的对称与否。将介绍的几种方法归纳列入表 9—27。当列联表的数据是从某一总体随机抽样获得时，若以样本推断两个变量在总体中是否相关，应采用表中最后一列的检验方法进行判定。

表 9--27　　两个变量的 PRE 测量方法

测量层次	相关量数	取值范围	假设检验
定类—定类	$\lambda(\lambda_{yx})$	[0, 1]	χ^2 检验
定序—定类	τ_y	[0, 1]	
定序--定序	G	[−1, +1]	Z 检验
	d_{yx}	[−1, +1]	
定类—定距 定序—定距	E^2	[0, 1]	F 检验

第十章　对数线性模型

列联表能够反映定类变量间较为复杂的关系，相应的统计检验及相关测量法，能够用来研究变量间的真实相关性和程度。但是，定类或定序变量无法建立如定距变量那样的回归模型、方差分析模型，这对描述变量间关系是很不方便的。对数线性模型是解决这一问题的极为有效的方法。虽然按照有无实参数的定义来划分，这种方法不能算作非参数统计方法，因为模型有参数估计问题。但它对定类数据，也包括定序数据建立模型很有效。受篇幅限制本书仅简介对数线性模型，以使读者能更好地了解、掌握和运用这一方法。

§10.1　高维列联表

当研究两个定类变量时，采用的是二维列联表。但许多问题的研究，往往涉及三个或三个以上的变量。如居民对电风扇的需求型式，可能与居住地区、家庭结构等有关。这时，二维列联表无法使用，而必须采用高维列联表。维是指变量，若涉及三个变量，则采用三维列联表，若涉及四个变量，则需用四维列联表，以此类推，可以得更高维的列联表。实际上，从三维表推广到四维或更高维表，除了增加分析的复杂性外，不存在更多的新问题，因而本书仅就三维表作一些分析。

一、三维表的命名法

二维列联表只有行、列，而三维列联表维数升高后，需增加“层”(Layer)，即具有行、列、层个类别。仍以 r、c 分别表示行、列，层用 l 表示，则三维列联表是 $r \times c \times l$ 表。例如，居民对电风扇的需求型式涉及三个变量：第一个是需求型式，可分为台式、台地式；第二个是居民居住地区，可分为热带、温带、寒带；第三个是家庭结

构，分为两代人家庭、三代及三代人以上家庭。调查结果如表10—1。表中每格也称作每单元的观察次数记作$f_{ijk}(i=1,2,\cdots,r;$ $j=1,2,\cdots,c;k=1,2,\cdots,l)$，通过对不同下标求和，可以得到各个边缘次数的总和。

表 10—1　　居民的电风扇需求型式　　单位：万台

		居住地区 (k)						合计
		热带		温带		寒带		
家庭结构(j)		两代	三代以上	两代	三代以上	两代	三代以上	
需求型式	台式	18	16	15	14	8	6	77
(i)	台地式	25	28	30	28	12	5	138
合计		43	44	45	42	20	21	215

单个变量边缘总次数。在i、j上对所有f_{ijk}值求和，得到第k层类的次数总和；若在j、k上对f_{ijk}求和，得到第i行类的次数总和；若在i、k上对f_{ijk}求和，得到第j列类的次数总和。其计算公式如(10.1)式。这是对两个下标求和的结果。利用表10—1数据可以计算得到

$$f_{i\cdot\cdot}=\sum_{j=1}^{c}\sum_{k=1}^{l}f_{ijk}$$

$$f_{\cdot j\cdot}=\sum_{i=1}^{r}\sum_{k=1}^{l}f_{ijk} \qquad (10.1)$$

$$f_{\cdot\cdot k}=\sum_{i=1}^{r}\sum_{j=1}^{c}f_{ijk}$$

$$f_{1\cdot\cdot}=(18+16)+(15+14)+(8+6)=77$$

$$f_{\cdot1\cdot}=(18+25)+(15+30)+(8+12)=108$$

$$f_{\cdot\cdot1}=18+16+25-28=87$$

$f_{1\cdot\cdot}$表示需求台式电风扇的数目，$f_{\cdot1\cdot}$表示家庭为两代人结构的电风扇需求数目，$f_{\cdot\cdot1}$表示热带居民电风扇的需求数目。类似地，根据表10—1还可以计算

台地式电风扇需求数目

$$f_{2\cdot\cdot}=(25+28)+(30+28)+(12+15)=138$$

三代人及三代人以上家庭电风扇需求数目

$$f_{\cdot 2\cdot} = (16+28)+(14+28)+(6+15)=107$$

温带地区居民电风扇需求数目

$$f_{\cdot\cdot 2} = 15+14+30+28=87$$

寒带地区居民电风扇需求数目

$$f_{\cdot\cdot 3} = 8+12+6+15=41$$

两个变量边缘总次数。对单个下标求和，可以得到两个变量边缘总次数。其计算公式如(10.2)式。

$$f_{ij\cdot} = \sum_{k=1}^{l} f_{ijk}$$

$$f_{i\cdot k} = \sum_{j=1}^{c} f_{ijk} \qquad (10.2)$$

$$f_{\cdot jk} = \sum_{i=1}^{r} f_{ijk}$$

表10—2是根据表10—1数据，通过对第三个变量，即下标 k 上求和得到的。

表10—2　　家庭结构和需求型式的边缘总次数

需求型式	家庭结构	
	两代	三代及以上
台式	41	36
台地式	67	71

其中

$$f_{11\cdot} = 18+15+8=41$$

$$f_{12\cdot} = 16+14+6=36$$

$$f_{21\cdot} = 25+30+12=67$$

$$f_{22\cdot} = 28+28+15=71$$

类似地，可以通过对第一个变量，即下标 i 上求和，得到第二、三个变量的边缘总次数；对第二个变量即下标 j 上求和，得到第一、三个变量的边缘总次数。同样地可以绘制出如上的表。

列联表总次数。通过对三个变量求和，可以得到列联表的总次数 n。其计算公式如(10.3)式。

$$n = f_{\cdots} = \sum_{i=1}^{r} \sum_{j=1}^{c} \sum_{k=1}^{l} f_{ijk} \qquad (10.3)$$

二、独立性检验

在二维列联表中，由于只涉及两个变量，因而独立性检验的假设也仅涉及两个变量之间。高维列联表中涉及的变量多于两个，因而可能研究的是所有变量间是否相互独立，也可能希望检验某些变量与其它一些变量相互独立，或是某一特定变量与其余变量无关。为此，高维表变量间独立性检验的假设，比二维表情况复杂。

1. 三个变量相互独立性的检验

若三个变量分别记作 X、Y、Z，则检验三个变量相互独立的假设组为

H_0:　X、Y、Z 间相互独立

H_1:　X、Y、Z 间不完全独立

为对假设作出判定，与二维表独立性检验相同，首先需要确定 H_0 为真时的检验统计量，然后得到其抽样分布，再确定 P 值，以 H_0 为真时，检验统计量为某值的概率 P 的大小作出拒绝或不拒绝 H_0 的判定。

设某一观察值出现在第 ijk 格中的概率为 $p_{ijk}(i = 1,2,\cdots,r; j = 1,2,\cdots,c; k = 1,2,\cdots,l)$ 第 i 行、第 j 列，第 k 层变量的边缘概率分别为 $p_{i\cdot\cdot}, p_{\cdot j\cdot}, p_{\cdot\cdot k}$，则 X、Y、Z 相互独立时有

$$p_{ijk} = p_{i\cdot\cdot}p_{\cdot j\cdot}p_{\cdot\cdot k}$$

因此，原假设 H_0 为

$$H_0: \quad p_{ijk} = p_{i\cdot\cdot}p_{\cdot j\cdot}p_{\cdot\cdot k}$$

当 H_0 为真时，即 $p_{ijk} = p_{i\cdot\cdot}p_{\cdot j\cdot}p_{\cdot\cdot k}$ 成立时，第 ijk 格的期望次数为

$$e_{ijk} = np_{i\cdot\cdot}p_{\cdot j\cdot}p_{\cdot\cdot k}$$

可以证明，$p_{i\cdot\cdot}$、$p_{\cdot j\cdot}$、$p_{\cdot\cdot k}$ 的最好估计值是利用相应的单个变量的边缘次数计算的频率，即

$$p_{i\cdot\cdot} = \frac{f_{i\cdot\cdot}}{n} \qquad p_{\cdot j\cdot} = \frac{f_{\cdot j\cdot}}{n} \qquad p_{\cdot\cdot k} = \frac{f_{\cdot\cdot k}}{n}$$

于是，期望次数 e_{ijk} 为

$$e_{ijk} = n \cdot \frac{f_{i..}}{n} \cdot \frac{f_{.j.}}{n} \cdot \frac{f_{..k}}{n}$$
$$= f_{i..}f_{.j.}f_{..k}/n^2 \qquad (10.4)$$

实际次数 f_{ijk} 与期望次数 e_{ijk} 越接近，表明变量间相互独立的可能性愈大，反之，f_{ijk} 与 e_{ijk} 的差值越大，表明变量间相互独立的可能性愈小。检验统计量 Q 是反映 f_{ijk} 与 e_{ijk} 差值大小的一个很好选择，这与二维表的独立性检验一样。Q 的计算如(10.5)式。

$$Q = \sum_{i=1}^{r} \sum_{j=1}^{c} \sum_{k=1}^{l} \frac{(f_{ijk} - e_{ijk})^2}{e_{ijk}} \qquad (10.5)$$

高维列联表情况下，统计量 Q 的自由度可以由(10.6)式计算得到。

df ＝(表的格数 － 1)

－(为检验特定假设须估计的概率数目)　　(10.6)

在三维表中，$r \times c \times l$ 的格数为 rcl。为检验 H_0，需要估计概率 $p_{i..}$，$p_{.j.}$ 和 $p_{..k}$。对 $p_{i..}$ 来说要估计 r 个，但由于概率和为 1，因此，只需估计 $(r-1)$ 个即可。同理，对于 $p_{.j.}$ 来说，需估计 $(c-1)$ 个，而 $p_{..k}$ 需估计 $(l-1)$ 个。对于 H_0 为真来说，统计量 Q 的自由度为

$$\mathrm{df} = (rcl - 1) - (r - 1) - (c - 1) - (k - 1)$$
$$= rcl - r - c - l + 2$$

确定 P 值。在附表 I 中，根据df，可以查找 H_0 为真时，Q 为某值的概率 P，将 P 与显著性水平 α 比较，若 P 足够小，则拒绝 H_0；若 P 相对 α 较大，则不能拒绝 H_0。与二维表的独立性检验一样，也可以在附表 I 中，根据 df 和显著性水平 α，查找临界值 χ^2_α，得到否定域。若 $Q \geqslant \chi^2_\alpha$，则拒绝 H_0；若 $Q < \chi^2_\alpha$，则不能拒绝 H_0。

【例 10.1】　利用表 10—1 的调查结果，分析说明居民居住地区、家庭结构与电风扇需求型式间是否相互独立

分析：设居民家庭结构、电风扇需求型式、居住地区分别为变量 X、Y、Z，研究三者间是否相互独立，是对所有变量之间相互独立性进行检验，建立原假设为

$$H_0: \quad p_{ijk} = p_{i\cdot\cdot}p_{\cdot j\cdot}p_{\cdot\cdot k}$$

即 X、Y、Z 间相互独立

根据表 10—1 和(10.1) 式及(10.4) 式，可以计算得到各个期望次数 e_{ijk}，如表 10—3。其中：

$$e_{111} = f_{1\cdot\cdot}f_{\cdot1\cdot}f_{\cdot\cdot1}/n^2 = 77(108)(87)/215^2 = 15.65,$$

$$e_{211} = f_{2\cdot\cdot}f_{\cdot1\cdot}f_{\cdot\cdot1}/n^2 = 138(108)(87)/215^2 = 28.05$$

其余期望次数的计算类似。由表 10—1 和表 10—3 可以得到检验统计量 Q。

表 10—3　　三个变量相互独立假设下的期望次数

		居住地区 (k)						合计
		热带		温带		寒带		
家庭结构(j)		两代	三代以上	两代	三代以上	两代	三代以上	
需求型式	台式	15.65	15.51	15.65	15.51	7.38	7.30	77
(i)	台地式	28.05	27.79	28.05	27.79	13.22	13.10	138
合计		43.7	43.3	43.7	43.3	20.6	20.4	215

$$Q = \sum_{i=1}^{r}\sum_{j=1}^{c}\sum_{k=1}^{l}\frac{(f_{ijk} - e_{ijk})^2}{e_{ijk}}$$

$$= \frac{(18 - 15.65)^2}{15.65} + \frac{(16 - 15.51)^2}{15.51} + \cdots + \frac{(15 - 13.10)^2}{13.10}$$

$$= 1.6846$$

查附表 Ⅰ，$\alpha = 0.05$，$df = rcl - r - c - l + 2 = 7$，得到临界值 $\chi^2_{0.05} = 14.07$。由于 $Q = 1.6846 < \chi^2_{0.05} = 14.07$，数据不能拒绝 H_0，表明在 5% 的显著性水平上，三个变量相互独立。

实际运算时，统计量 Q 的计算公式还可以采用另一种形式，即(10.7) 式。

$$Q = n^2\left[\sum_{i=1}^{r}\sum_{j=1}^{c}\sum_{k=1}^{l}\frac{f_{ijk}^2}{f_{i\cdot\cdot}f_{\cdot j\cdot}f_{\cdot\cdot k}} - \frac{1}{n}\right] \qquad (10.7)$$

(10.7) 式也可以写成

$$Q = n^2 \sum_{i=1}^{r} \sum_{j=1}^{c} \sum_{k=1}^{l} \frac{f_{ijk}^2}{f_{i..} f_{.j.} f_{..k}} - n \text{ ①} \qquad (10.8)$$

例 10.1 中三个变量相互独立，因此不必要再作进一步的分析。

【例 10.2】 呼吸情况与年龄、吸烟状况是否有关

某地区随机抽取 97 人，对他们的呼吸情况、年龄和吸烟状况进行调查，结果如表 10—4。

分析：为了检验这批人的呼吸情况与年龄、吸烟状况是否有关，可以建立假设组

H_0： 呼吸情况、年龄、吸烟状况相互独立

H_1： 呼吸情况、年龄、吸烟状况不相互独立

计算检验统计量 Q，采用 χ^2 检验对假设作出判定。统计量 Q 可以根据(10.5)式或(10.6)计算。为计算简便，采用(10.6)式。由于

表 10—4 呼吸情况与年龄、吸烟状况调查表

年龄(i)	吸烟状况(j)	呼吸情况(k)			合计
		正常	尚可	异常	
<40	从不吸烟	16	15	5	36
	吸烟	7	34	3	44
40—59	从不吸烟	1	3	1	5
	吸烟	1	8	3	12
合计		25	60	12	97

$$f_{1..} = (16 + 15 + 5) + (7 + 34 + 3) = 80$$

$$f_{2..} = (1 + 3 + 1) + (1 + 8 + 1) = 17$$

$$f_{.1.} = (16 + 15 + 5) + (1 + 3 + 1) = 41$$

① $Q = \sum_{i=1}^{r} \sum_{j=1}^{c} \sum_{k=1}^{l} \frac{(f_{ijk} - e_{ijk})^2}{e_{ijk}} = \Sigma\Sigma\Sigma \frac{f_{ijk}^2 - 2f_{ijk}e_{ijk} + e_{ijk}^2}{e_{ijk}}$

$= \Sigma\Sigma\Sigma \frac{f_{ijk}^2}{e_{ijk}} - 2\Sigma\Sigma\Sigma f_{ijk} + \Sigma\Sigma\Sigma e_{ijk} = n^2 \Sigma\Sigma\Sigma \frac{f_{ijk}}{f_{i..} f_{.j.} f_{..k}} - n$

$$f_{\cdot 2\cdot} = (7 + 34 + 3) + (1 + 8 + 3) = 56$$

$$f_{\cdot\cdot 1} = 16 + 7 + 1 + 1 = 25$$

$$f_{\cdot\cdot 2} = 15 + 34 + 3 + 8 = 60$$

$$f_{\cdot\cdot 3} = 5 + 3 + 1 + 3 = 12$$

所以

$$Q = n^2 \sum_{i=1}^{r} \sum_{j=1}^{c} \sum_{k=1}^{l} \frac{f_{ijk}}{f_{i\cdot\cdot} f_{\cdot j\cdot} f_{\cdot\cdot k}} - n = 15.908$$

在附表 I 中，$\alpha = 0.05$，df $= 7$，查得临界值为 $\chi^2_{0.05} = 14.07$。因为 $Q = 15.908 > \chi^2_{0.05} = 14.06$，所以数据在 5% 的显著性水平上拒绝 H_0。这表明三个变量不是相互独立的。

当调查数据拒绝 H_0 时，可以对列联表作进一步的分析，研究是由于哪些变量引起拒绝 H_0。

2. 局部独立性检验

三个变量间相互独立性的假设被拒绝，并不意味着所有变量之间都存在着显著的联系。可能是两个变量间相关，而第三个变量完全独立，即有局部独立性；也可能是两个变量在第三个变量的每一水平上是独立的，但两个变量的每一个都与第三个变量相关，即当给定第三个变量的水平时，前两个变量是条件独立的。为了能更深入地研究变量间的关系，常常需要对列联表作深入的分析。例如，对三维表建立下述的三个独立性假设：

$H_0^{(1)}$：　$p_{ijk} = p_{i\cdot\cdot} p_{\cdot jk}$　（行分类独立于列和层分类）

$H_0^{(2)}$：　$p_{ijk} = p_{\cdot j\cdot} p_{i\cdot k}$　（列分类独立于行和层分类）

$H_0^{(3)}$：　$p_{ijk} = p_{\cdot\cdot k} p_{ij\cdot}$　（层分类独立于列和行分类）

对于 $H_0^{(1)}$ 来说，$p_{ijk} = p_{i\cdot\cdot} p_{\cdot jk}$ 成立，意味着，$p_{ij\cdot} = p_{i\cdot\cdot} p_{\cdot j\cdot}$ 并且 $p_{i\cdot k} = p_{i\cdot\cdot} p_{\cdot\cdot k}$ 成立，这也就是说，$H_0^{(1)}$ 是行分类和列分类独立与行分类和层分类独立的混合假设。

为检验假设，可以按照与以前相同的方式进行。若 $H_0^{(1)}$ 为真，即 $p_{ijk} = p_{i\cdot\cdot} p_{\cdot jk}$ 成立时，第 ijk 格的期望次数为

$$e_{ijk} = n p_{i\cdot\cdot} p_{\cdot jk}$$

可以证明，概率 $p_{i\cdot\cdot}$、$p_{\cdot jk}$ 的最好估计值是频率即有

$$p_{i\cdot\cdot}=\frac{f_{i\cdot\cdot}}{n}\qquad p_{\cdot jk}=\frac{f_{\cdot jk}}{n}$$

因此

$$e_{ijk}=n\cdot\frac{f_{i\cdot\cdot}}{n}\cdot\frac{f_{\cdot jk}}{n}=\frac{f_{i\cdot\cdot}f_{\cdot jk}}{n}\tag{10.9}$$

实际次数 f_{ijk} 与期望次数 e_{ijk} 的差值，可以用来判定变量间相互独立的可能性。其差值越小，相互独立的可能性越大，反之，差值越大，相互独立的可能性越小。检验统计量仍为 Q，其计算公式为

$$Q=\sum_{i=1}^{r}\sum_{j=1}^{c}\sum_{k=1}^{l}\frac{(f_{ijk}-e_{ijk})^2}{e_{ijk}}$$

这与(10.5)式相同。Q 统计量遵从 χ^2 分布，其自由度仍用(10.6)式确定。表的格数为 rcl，为检验假设要估计 $p_{i\cdot\cdot}$ 和 $p_{\cdot jk}$，估计 $p_{i\cdot\cdot}$ 的数目为 $(r-1)$ 个，而估计 $p_{\cdot jk}$ 的数目即列 × 层的概率数目为 $(cl-1)$ 个。因此，自由度为

$$\mathrm{df}=(rcl-1)-(r-1)-(cl-1)=rcl-r-cl+1$$

χ^2 检验中确定 P 值的方法在这里同样适用。根据df、显著性水平 α，以及 Q 值，可以对假设作出判定。

【例 10.3】 年龄与吸烟状况和呼吸情况是否无关

在例 10.2 中，年龄、吸烟状况、呼吸情况这三个变量被证明不是相互独立的，但并不一定说明三个变量间都有显著联系。本例要利用表 10—4 的数据检验年龄是否独立于其他两个变量。

分析：根据题目的要求，应建立假设组

$H_0^{(1)}$： $p_{ijk}=p_{i\cdot\cdot}p_{\cdot jk}$

$H_1^{(1)}$： $p_{ijk}\neq p_{i\cdot\cdot}p_{\cdot jk}$

用文字表述为

$H_0^{(1)}$： 年龄与吸烟状况和呼吸情况无关

$H_1^{(1)}$： 年龄与吸烟状况和呼吸情况有关

或

$H_0^{(1)}$： 年龄独立于吸烟状况和呼吸情况

$H_1^{(1)}$： 年龄不独立于吸烟状况和呼吸情况

根据表 10—4 的数据，按照(10.9)式计算各个期望次数，如表 10—5。

表 10—5　在年龄和其它两个变量无关假设下的期望次数

年龄(i)	吸烟状况(j)	呼吸情况(k)			合计
		正常	尚可	异常	
<40	从不吸烟	14.02	14.84	4.95	33.81
	吸烟	6.60	34.64	4.95	46.19
40—59	从不吸烟	2.98	3.16	1.05	7.19
	吸烟	1.40	7.36	1.05	9.81
合计		25.00	60.00	12.00	97

由表 10—4 可知

$$f_{\cdot 11} = 16 + 1 = 17 \qquad f_{\cdot 12} = 15 + 3 = 18$$

$$f_{\cdot 21} = 7 + 1 = 8 \qquad f_{\cdot 22} = 34 + 8 = 42$$

$$f_{\cdot 13} = 5 + 1 = 6 \qquad f_{\cdot 23} = 3 + 3 = 6$$

因而有

$$e_{111} = (f_{1\cdot\cdot} f_{\cdot 11})/n = 80(17)/97 = 14.02$$

$$e_{112} = (f_{1\cdot\cdot} f_{\cdot 12})/n = 80(18)/97 = 14.84$$

以此类推，得到表 10—5 的各个期望次数。于是

$$Q = \sum_{i=1}^{r}\sum_{j=1}^{c}\sum_{k=1}^{l}\frac{(f_{ijk} - e_{ijk})^2}{e_{ijk}}$$

$$= \frac{(16 - 14.02)^2}{14.02} + \cdots + \frac{(3 - 1.05)^2}{1.05}$$

$$= 6.2035$$

在附表 I 中，根据 $df = rcl - r - cl + 1 = 5$，显著性水平 $\alpha = 0.05$，查得 $\chi^2_{0.05} = 11.07$。由于 $Q = 6.2035 < \chi^2_{0.05} = 11.07$，所以数据在 5% 的显著性水平上不能拒绝 H_0，表明年龄独立于其他两个变量。

由例 10.2 和例 10.3 的结论可以得知，吸烟状况与呼吸情况有关。利用表 10—4 的数据可以进一步验证这一结论。对表10—4的数据在年龄变量上求和，即对 i 求和，可以将 $2\times2\times3$ 列联表折迭成一张 2×3 列联表，如表 10—6。

表 10—6　表 10—4 数据对年龄变量求和的结果

吸烟状况	呼吸情况			合计
	正常	尚可	异常	
从不吸烟	17(10.57)	18(25.36)	6(5.07)	41
吸烟	8(14.43)	42(34.64)	6(6.93)	56
合计	25	60	12	97

根据检验目的建立的假设组为

H_0:　吸烟状况与呼吸情况无关

H_1:　吸烟状况与呼吸情况相关

按照 H_0 为真时，二维列联表各个期望次数的计算公式，可以得到独立性假设下的各个期望次数，如表 10—6 中括号内的值，利用(9.4)式计算 Q 统计量得到

$$Q=97\left[\frac{17^2}{25(41)}+\cdots+\frac{6^2}{12(56)}-1\right]$$
$$=10.7864$$

在附表 I 中，根据 $df=(r-1)(c-1)=2$，显著性水平 $\alpha=0.01$，查得 $\chi^2_{0.01}=9.21$。因为 $Q=10.7864>\chi^2_{0.01}=9.21$，所以数据在 1% 的显著性水平上拒绝 H_0，表明吸烟状况确实与呼吸情况相关。对于其他的假设，如 $H_0^{(2)}$、$H_0^{(3)}$ 等，可以采用类似的方法，得到各单元的期望值 e_{ijk}。将实际次数 f_{ijk} 与期望次数 e_{ijk} 代入(10.5)式，得到对应于某个假设条件下的统计量 Q。按照(10.6)式确定自由度后，能够对假设作出判定。

§10.2　对数线性模型的类型和参数估计

前面关于列联表分析的假设检验，能够有助于认识变量间的

较为复杂的关系，但还没有将这些关系量化，即没有建立起关于变量间关系的模型，对数线性模型恰恰解决了这一问题。

一、引言

在二维列联表中，检验两个变量的相互独立性建立的原假设为

$$p_{ij} = p_{i\cdot}p_{\cdot j} \tag{10.10}$$

这一等式表明，在总体内，一次观察落入表中第 ij 格的概率为边缘概率的乘积，这实际上也就确定了数据的结构或模型。对(10.10)式两边同时取对数得到

$$\ln p_{ij} = \ln p_{i\cdot} + \ln p_{\cdot j} \tag{10.11}$$

由于原假设成立时，期望次数 $e_{ij} = n \cdot p_{ij}$，因此(10.11)式可以写成

$$\ln e_{ij} = \ln e_{i\cdot} + \ln e_{\cdot j} - \ln n \tag{10.12}$$

将(10.12)式两边同时对 i 求和有

$$\sum_{i=1}^{r} \ln e_{ij} = \sum_{i=1}^{r} \ln e_{i\cdot} + r\ln e_{\cdot j} - r\ln n$$

将(10.12)式两边同时对 j 求和有

$$\sum_{j=1}^{c} \ln e_{ij} = c\ln e_{i\cdot} + \sum_{j=1}^{c} \ln e_{\cdot j} - c\ln n$$

将(10.12)式两边同时对 i、j 求和有

$$\sum_{i=1}^{r}\sum_{j=1}^{c} \ln e_{ij} = c\sum_{i=1}^{r} \ln e_{i\cdot} + r\sum_{j=1}^{c} \ln e_{\cdot j} - rc\ln n$$

若定义

$$\lambda = \frac{\sum_{i=1}^{r}\sum_{j=1}^{c} \ln e_{ij}}{rc} \tag{10.13}$$

$$\alpha_i = \frac{\sum_{i=1}^{r} \ln e_{ij}}{c} - \frac{\sum_{i=1}^{r}\sum_{j=1}^{c} \ln e_{ij}}{rc} \tag{10.14}$$

$$\beta_j = \frac{\sum_{i=1}^{r} \ln e_{ij}}{c} - \frac{\sum_{i=1}^{r}\sum_{j=1}^{c} \ln e_{ij}}{rc} \tag{10.15}$$

那么,(10.12)式可以写成

$$\ln e_{ij} = \lambda + \alpha_i + \beta_j \quad (i = 1,2,\cdots,r, j = 1,2,\cdots,c) \quad (10.16)$$

这类似于方差分析中所用的模型。借助于方差分析术语。(10.16)式中,λ 表示"总平均效应",α_i 表示行变量第 i 类的"主效应",β_j 表示列变量第 j 类的"主效应"。由(10.14)式和(10.15)式可知,α_i 测度了频数对数的行平均与总平均间的偏差,而 β_j 测度了频数对数的列平均与总平均间的偏差,因此有

$$\sum_{i=1}^{r} \alpha_i = 0 \qquad \sum_{j=1}^{c} \beta_j = 0$$

(10.16)式是二维列联表两个变量相互独立时的对数线性模型。

二、对数线性模型的类型

对数线性模型是用来分析定类数据的一类很有用的数学模型,可以用于反映列联表中各个变量间的复杂关系。由(10.16)式可知,对数线性模型是将列联表上每单元的频数作为因变量,表上所有变量作为自变量,建立各个自变量的效应与每单元频数的对数之间的函数关系,用以分析研究表上各个变量间的关系。(10.16)式是两个变量相互独立时的对数线性模型。事实上,变量间可能并不相互独立,当变量多于两个时,变量间的关系还会更复杂。因此,对数线性模型有很多类型。本章只讨论几种常用的且较为简单的类型。

(一)饱和模型(Saturated Model)

当变量间相互不独立时,变量间的相关会使变量相互作用,可以用方差分析中的一个术语"交互作用(interaction)"来描述列联表中变量间的这种关系。变量间相互不独立时的对数线性模型称为饱和模型。饱和模型中,不仅有各个变量的主效应,还应有变量间的交互作用效应。

1. 两个变量的饱和模型

对于两个变量 X、Y 的 $r \times c$ 列联表,饱和模型为

相互不独立，因此 λ_{ij}^{12} 不为 0。模型 Ⅱ 类是三个变量的条件独立性关系，如三个变量虽然相互不独立，但给定变量 Z 时，X 与 Y 是独立的，因此 $\lambda_{ij}^{12}=0$，当然有 $\lambda_{ijk}^{123}=0$。根据谱系模型的规则，诸如模型 $\lambda+\lambda_i^1+\lambda_j^2+\lambda_k^3+\lambda_{ijk}^{123}$ 是不允许的。这一点应予以注意。

三、模型参数的估计

对数线性模型的一个主要优点是能够估计模型中各个参数，而这些参数的值，使各个变量的效应和变量间的交互作用效应得以数量化。

(一) 直接计算

在某些情况下，对数线性模型可以利用列联表的各边缘次数直接计算各个效应参数。

两个变量的饱和模型，如(10.17)式，其参数可以利用列联表的各个频数直接计算得到。以 $\mu_{..}$ 表示各个观察值对数的总平均值，即总平均效应，则有 $\mu_{..}=\lambda=(\Sigma\Sigma\ln f_{ij})/rc$；以 $\mu_{i.}$、$\mu_{.j}$ 分别表示第 i 行、第 j 列频数对数的平均值，即有

$$\mu_{i.}=(\sum_{j=1}^{c}\ln f_{ij})/c$$

$$\mu_{.j}=(\sum_{i=1}^{r}\ln f_{ij})/r$$

那么，模型的各个参数值计算公式为

$$\lambda_i^1=\mu_{i.}-\mu_{..} \qquad (10.21)$$

$$\lambda_j^2=\mu_{.j}-\mu_{..} \qquad (10.22)$$

$$\begin{aligned}\lambda_{ij}^{12}&=\ln f_{ij}-(\mu_{..}+\lambda_i^1+\lambda_j^2)\\&=\ln f_{ij}-\mu_{i.}-\mu_{.j}+\mu_{..}\end{aligned} \qquad (10.23)$$

主效应系数 λ_i^1 或 λ_j^2 若大于 0，表明效应为正；若小于 0，表明效应为负。λ_i^1 是第一个变量的第 i 个水平对总平均效应 λ 的增减量；λ_j^2 是第二个变量的第 j 个水平对总平均效应 λ 的增减量。λ_{ij}^{12} 代表变量 1 和变量 2 在各自的第 i 个水平和第 j 个水平之间交互作用效应，是其交互作用对总平均效应 λ 的增减量。λ_{ij}^{12} 大于 0，表明效应为正；若

λ_{ij}^{12} 小于 0，则效应为负。

【例 10.4】 居民居住地区与电风扇需求型式的调查结果如表 9—7。若满足(10.17)式的关系，试估计各效应参数。

表 10—9　　　表 9—7 各单元频数的对数值

电风扇需求型式（变量 1）	居住地区（变量 2）			均值
	热带	温带	寒带	
台式	2.6391	3.4012	1.3863	2.4755
落地式	4.2047	4.6540	4.0943	4.3177
台地式	3.4012	2.5649	2.6391	2.8684
均　值	3.4150	3.5400	2.7060	3.2205

解：根据表 9—7 的数据，计算表上各单元频数的自然对数，列入表 10—9，同时计算该表各行、各列的平均值及总平均值。根据(10.21)式有

$$\lambda_1^1 = \mu_{1\cdot} - \mu_{\cdot\cdot}$$
$$= 2.4755 - 3.2205 = -0.745$$
$$\lambda_2^1 = \mu_{2\cdot} - \mu_{\cdot\cdot}$$
$$= 4.3177 - 3.2205 = 1.097$$

由于 $\sum_{i=1}^{3} \lambda_i^1 = 0$，所以

$$\lambda_3^1 = -(\lambda_1^1 + \lambda_2^1) = -(-0.745 + 1.097) = -0.352$$

根据(10.22)式有

$$\lambda_1^2 = \mu_{\cdot 1} - \mu_{\cdot\cdot} = 3.4150 - 3.2205 = 0.195$$
$$\lambda_2^2 = \mu_{\cdot 2} - \mu_{\cdot\cdot} = 3.5400 - 3.2205 = 0.320$$

由于 $\sum_{j=1}^{3} \lambda_j^2 = 0$，所以，$\lambda_3^2 = -(\lambda_1^2 + \lambda_2^2) = -(0.195 + 0.320) = -0.515$

各个主效应的估计值如表 10—10。

表 10—10　　　　**主效应的估计值**

		变量 1	变量 2
水平	1	$\lambda_1^1 = -0.745$	$\lambda_1^2 = 0.195$
	2	$\lambda_2^1 = 1.097$	$\lambda_2^2 = 0.320$
	3	$\lambda_3^1 = -0.352$	$\lambda_3^2 = -0.515$

根据(10.23)式有

$$\lambda_{11}^{12} = \ln f_{11} - \mu_{1\cdot} - \mu_{\cdot 1} + \mu_{\cdot\cdot}$$
$$= 2.6391 - 2.4755 - 3.4150 + 3.2205 = -0.031$$
$$\lambda_{21}^{12} = \ln f_{21} - \mu_{2\cdot} - \mu_{\cdot 1} + \mu_{\cdot\cdot} = -0.308$$

由于 $\sum_{i=1}^{3}\lambda_{ij}^{12} = 0$，所以，$\lambda_{31}^{12} = -(\lambda_{11}^{12} + \lambda_{21}^{12}) = -(-0.031 - 0.308) = 0.399$。同理，根据(10.23)式计算出 λ_{21}^{12}、λ_{22}^{12} 后，可以利用 $\sum_{i=1}^{3}\lambda_{ij}^{12} = 0$，推算 λ_{32}^{12}。又由于 $\sum_{j=1}^{3}\lambda_{ij} = 0$，因而利用前面计算得到的 λ_{11}^{12}、λ_{12}^{12} 可以推算出 λ_{13}^{12}，同样的方法能够推算出 λ_{23}^{12}、λ_{33}^{12}。将各交互效应估计值列入表 10—11。

表 10—11　　**变量间交互作用效应估计值**

$\lambda_{11}^{12} = -0.031$	$\lambda_{21}^{12} = -0.308$	$\lambda_{31}^{12} = 0.339$
$\lambda_{12}^{12} = 0.606$	$\lambda_{22}^{12} = 0.017$	$\lambda_{32}^{12} = -0.623$
$\lambda_{13}^{12} = -0.575$	$\lambda_{23}^{12} = 0.291$	$\lambda_{33}^{12} = 0.284$

从上面参数计算过程可以看出，饱和模型需估计的参数数目恰恰与列联表单元数目相等，因而模型对数据完全拟合。例 10.4 中列联表单元数目为 $3 \times 3 = 9$，而估计的参数为 λ、λ_1^1、λ_2^1、λ_1^2、λ_2^2、λ_{11}^{12}、λ_{12}^{12}、λ_{21}^{12}、λ_{22}^{12}、，也是 9 个，其余参数均可由关系式导出。根据表 10—10，表 10—11，有

$$\ln f_{22} = \lambda + \lambda_2^1 + \lambda_2^2 + \lambda_{22}^{12}$$
$$= 3.220 + 1.097 + 0.320 + 0.017 = 4.654$$

这一结果表明，对于居住温带地区且需求落地式电风扇来说，两个

变量各自主效应以及两个变量的交互作用效应均为正效应。由表10—10可知，台式、台地式电风扇需求是负效应，对总平均效应 λ 起减少作用，而落地式电风扇需求是正效应，起增加作用；变量2中，居住热带、温带均为正效应，居住寒带为负效应。由表10—11则可进一步看出两个变量交互作用的结果。落地式电风扇需求效应为正效应，但其与居住地区交互作用时，仅与温带、寒带交互作用产生正效应，而与热带交互作用时是负效应。

【例10.5】 在例10.2中，年龄、吸烟情况和呼吸状况相互不独立，但在例10.3中经检验发现年龄独立于其它两个变量，即三个变量间有如表10—8中类型 Ⅱ 的模型关系：$\lambda+\lambda_i^1+\lambda_j^2+\lambda_k^3+\lambda_{jk}^{23}$。试估计各个效应参数。

解：假定各个效应参数满足下列关系

$$\sum_{i=1}^{r}\lambda_i^1=0,\quad \sum_{j=1}^{c}\lambda_j^2=0,\quad \sum_{k=1}^{l}\lambda_k^3=0,\text{且}$$

$$\Sigma\lambda_{jk}^{23}=0,\quad \Sigma\lambda_{jk}^{23}=0$$

那么，也可以利用列联表频数对数的行、列、层边缘值直接计算各效应参数。

总平均效应参数 λ 为总平均值，因此有

$$\lambda=\mu_{\cdots}=(\sum_{i=1}^{r}\sum_{j=1}^{c}\sum_{k=1}^{l}\ln f_{ij})/rcl \tag{10.24}$$

λ_i^1 为变量1的主效应，反映变量1的平均效应与总平均效应的偏差，因此有

$$\lambda_i^1=\mu_{i\cdot\cdot}-\mu_{\cdots} \tag{10.25}$$

同理可以得到

$$\lambda_j^2=\mu_{\cdot j\cdot}-\mu_{\cdots} \tag{10.26}$$

$$\lambda_k^3=\mu_{\cdot\cdot k}-\mu_{\cdots} \tag{10.27}$$

λ_{jk}^{23} 是变量2和变量3的交互作用效应，因此有

$$\begin{aligned}\lambda_{jk}^{23}&=(\mu_{\cdot jk}-\mu_{\cdots}-\lambda_j^2-\lambda_k^3)\\&=\mu_{\cdot jk}-\mu_{\cdot j\cdot}-\mu_{\cdot\cdot k}+\mu_{\cdots}\end{aligned} \tag{10.28}$$

将表10—4中的数据取自然对数，得到表10—12。

0.0169。又由于 $\sum_{j=1}^{2} \lambda_{j\cdot}^{23} = 0$,所以其它交互作用效应参数能够得到。各个交互作用效应参数估计值列入表 10—14。

表 10—14　吸烟状况与呼吸情况交互作用效应的估计值

$\lambda_{11}^{23} = 0.3367$	$\lambda_{21}^{23} = -0.3367$
$\lambda_{12}^{23} = -0.3198$	$\lambda_{22}^{23} = 0.3198$
$\lambda_{13}^{23} = -0.0169$	$\lambda_{23}^{23} = 0.0169$

主效应大于 0,表明效应为正,如 $\lambda_1^1 = 0.7820 > 0$,是因为年龄在 40 岁以下的人比 40 岁以上的人多;主效应小于 0,表明效应为负,如 $\lambda_1^2 = -0.13 < 0$,是因为从不吸烟的人少于吸烟的人。交互效应大于 0,表明其交互作用效应为正,如 $\lambda_{11}^{23} = 0.3367 > 0$,表示从不吸烟而呼吸正常的人数,多于仅由从不吸烟的人数或呼吸正常的人数为基础估计的平均人数,也就是表明从不吸烟类与呼吸正常类之间存在着相关,因此这类人数比假定两个变量(吸烟状况与呼吸情况)无关时所期望的人数要多。交互效应小于 0,表明其交互作用效应为负,如 $\lambda_{21}^{23} = -0.3367 < 0$,表明吸烟而呼吸正常的人数少于仅由吸烟人数或呼吸正常人数为基础估计的平均人数,也就是说,吸烟类和呼吸正常类存在着负相关。

对数线性模型的各个效应系数是否显著不为零,也需要进行统计检验。这在后面将会介绍。非饱和模型与饱和模型不同,其不能对列联表的数据完全拟合。由表 10—13、表 10—14 可知,$\lambda + \lambda_1^1 + \lambda_1^2 + \lambda_1^3 + \lambda_{11}^{23} = 1.4948 + 0.7820 - 0.13 - 0.3152 + 0.3367 = 2.1683 \neq \ln f_{111} = 2.7726$。这是由于模型中需要估计的参数数目少于列联表上单元的数目。在例 10.5 中,列联表的单元数目为 $2 \times 2 \times 3 = 12$,而需要估计的参数只有 6 个,其余均由效应参数间关系推算而得。

用对数线性模型分析高维列联表,需要样本数目很大,因此,实际应用时,常常采用“压缩”的方法,将高维表就变量中的某一个折迭,成为低维列联表。一般来说,三个变量间是局部独立性关

系，即以表 10—8 中模型 Ⅱ 类拟合数据满意时，就三个变量中的任何一个折迭列联表，简化以后分析都是可行的；而三个变量是条件独立性关系，即需用模型 Ⅲ 类拟合数据时，以哪个变量折迭列联表，应持谨慎的态度，以避免导致虚假的结果。例如，例 10.3 表明，年龄与吸烟状况、呼吸情况是局部独立性关系，因而可以对年龄变量求和，将 2 × 2 × 3 表折迭成 2 × 3 表，如表 10—6。利用该表建立对数线性模型，研究两个变量的主效应以及交互作用效应。表 10—15 是表 10—6 中各频数的自然对数及均值。表 10—16 是各个主效应，表 10—17 是各个交互作用效应。

表 10—15　　与表 10—6 相对应的各频数对数

吸烟状况 （变量 1）	呼吸情况（变量 2）			均值
	正常	尚可	异常	
从不吸烟	2.8332	2.8904	1.7918	2.5051
吸　烟	2.0794	3.7377	1.7918	2.5363
均　值	2.4563	3.3140	1.7918	2.5207

表 10—16　　主效应估计值

		变量 1	变量 2
类	1	$\lambda_1^1 = -0.0156$	$\lambda_1^2 = -0.0644$
	2	$\lambda_2^1 = 0.0156$	$\lambda_2^2 = 0.7933$
	3		$\lambda_3^2 = -0.7289$

表 10—17　　交互作用效应估计值

$\lambda_{11}^{12} = 0.3925$	$\lambda_{21}^{12} = -0.3925$
$\lambda_{12}^{12} = -0.4080$	$\lambda_{22}^{12} = 0.4080$
$\lambda_{13}^{12} = 0.0155$	$\lambda_{23}^{12} = -0.0155$

对比表 10—16 和表 10—13 可以看出，无论是否考虑年龄变

量，吸烟状况的主效应作用方向一致。从不吸烟的主效应为负，吸烟的主效应为正，但对总平均效应 μ 的增减量不同。在不考虑年龄变量时，表 10—16 的结果显示从不吸烟对 μ 的减少量为 0.0156，而考虑年龄变量时，表 10—13 显示的减少量为 0.13。呼吸情况变量也有类似的结果。对比表 10—17 和表 10—14 可以看出，吸烟状况与呼吸情况两个变量交互作用效应受年龄变量的影响较大。从不吸烟与呼吸异常交互作用在考虑年龄时为负效应，而不考虑年龄时为正效应；吸烟与呼吸异常的交互作用恰恰相反，考虑年龄时为正效应，不考虑年龄时为负效应；并且交互作用对总平均效应 μ 的增减量都有所不同。

【例 10.6】 调查 A、B 两所医院产妇护理及婴儿生存情况的资料如表 10—18。是否能按 X 变量折迭列联表。

表 10—18 产前护理次数与婴儿生存情况

医院(X)	产前护理次数(Y)	婴儿生存情况(Z)	
		已 死	活 着
A	较少	3	176
	较多	4	293
B	较少	17	197
	较多	2	23

分析：根据表 10—18 的数据，分别考察 A、B 医院的产前护理次数与婴儿生存情况是否相关。利用 χ^2 检验判断。按(9.4)式计算 Q 统计量得到

$$Q_A = n[\Sigma\Sigma f_{ij}^2/f_{i\cdot}f_{\cdot j} - 1]$$
$$= 476[9/7(179) + 16/7(297) + 176^2/469(179)$$
$$+ 293^2/469(297) - 1]$$
$$= 0.0476$$
$$Q_B = n[\Sigma\Sigma f_{ij}^2/f_{i\cdot}f_{\cdot j} - 1]$$
$$= 239[17^2/19(214) + 4/19(25) + 197^2/220(214)$$

$$+ 23^2/220(25) - 1]$$
$$= 0.0000418$$

Q_A、Q_B均接近于0，表明数据不拒绝H_0，即在A医院、B医院，产前护理次数与婴儿生存情况相互独立。对于变量X、Y、Z来说，存在条件独立性，也就是在X的每个水平上，Y与Z相互独立。这种情况下，若对医院变量求和，将$2\times2\times2$表折迭成2×2表，如表10—19，考察产前护理次数与婴儿生存情况的相关情况。仍利用χ^2检验，计算Q统计量为

$$Q = n[\Sigma\Sigma f_{ij}^2/f_{i.}f_{.j} - 1]$$
$$= 715[20^2/26(393) + 373^2/689(393) + 6^2/26(322)$$
$$+ 316^2/689(332) - 1]$$
$$= 5.256$$

表10—19　产前护理次数与婴儿生存情况

产前护理次数	婴儿生存情况		合　计
	已死	活着	
较　少	20	373	393
较多	6	316	322
合　计	26	689	715

根据$df=(r-1)(c-1)=1$，$\alpha=0.05$；查附表Ⅰ，得到$\chi^2_{0.05}=3.84$。由于$Q=5.256>\chi^2_{0.05}=3.84$，所以，数据表明产前护理次数与婴儿生存情况相关。显然，这一结论是错误的。在三个变量具有条件独立性关系时，折迭列联表可能导致错误结论。这时，采用对数线性模型，考察变量间的关系是最合适的。

（二）迭代计算法

对于不能直接求解的模型，可以采用迭代法求解各个效应系数估计值。迭代计算法亦称重复估算法，用于独立模型和谱系模型的效应系数估计。迭代过程直到两次迭代估计值之差小于给定的收敛标准为止。一般采用极大似然估计准则。有兴趣的读者可参看

不确定性是较为适宜的。对 $-\ln p_i$ 按概率加权平均就得到随机变量 ζ 的熵，这就是(10.29)式。若随机变量 ζ 为连续的，其联合密度函数为 $f(x)$，则(10.29)式右端求和改为积分，即 ζ 的熵为

$$H(\zeta)=-\int f(x)\ln f(x)\mathrm{d}x \tag{10.30}$$

熵可以拓广到两个变量、三个变量甚至到更多个变量。若是考虑两个随机变量 X、Y，则它们的联合熵 $H(X,Y)$ 为

$$H(X,Y)=-\sum_{i=1}^{r}\sum_{j=1}^{c}p_{ij}\ln p_{ij} \tag{10.31}$$

若考察三个变量 X、Y、Z，则其联合熵为

$$H(X,Y,Z)=-\sum_{i=1}^{r}\sum_{j=1}^{c}\sum_{k=1}^{l}p_{ijk}\ln p_{ijk} \tag{10.32}$$

熵反映随机变量的不确定性。由其定义可知，熵的减少表明“信息”的增加。因此，变量间是否相关，可以用熵定量给出，当变量间相互独立时，联合的熵达到最大。

似然比是列联表中变量间相互独立时，相应的似然比函数的极大值与不独立时相应似然函数的极大值之比。似然比统计量常用来检验变量间的独立性。由于利用熵的减少可以研究变量间是否相关，即独立性，而用频率代替概率，直接根据列联表的频数可以计算得到反映变量相互独立时的熵，恰恰这也就是似然比统计量 Λ 的函数 $-2\ln\Lambda$。因此，检验对数线性模型拟合的效果，也就是检验变量间存在何种关系，可以采用熵或似然比。当样本量趋于无穷时，$-2\ln\Lambda$ 的极限分布是 χ^2 分布，所以常也称作 χ^2 似然比，亦译成 χ^2 拟合度。

2. 两个变量的 χ^2 似然比

当两个变量 X、Y 不独立时，联合熵为

$$H(X,Y)=-\sum_i\sum_j p_{ij}\ln p_{ij}$$

而当 X、Y 相互独立时，有 $p_{ij}=p_{i\cdot}p_{\cdot j}$，则联合熵为

$$H_{独立}(X,Y)=-\sum_i\sum_j p_{i\cdot}p_{\cdot j}\ln p_{i\cdot}p_{\cdot j}$$

X、Y 独立时，联合的熵达到最大，因此，可以用上述两式的差表示

X、Y 间的相关情况，即有

$$H_{独立}(X,Y) - H(X,Y)$$
$$= \sum_i \sum_j p_{ij} \ln p_{ij} - \sum_i \sum_j p_{i\cdot} p_{j\cdot} \ln p_{i\cdot} p_{\cdot j}$$

由于 $\Sigma \Sigma p_{i\cdot} p_{\cdot j} = 1$，所以上式可以写成

$$H_{独立}(X,Y) - H(X,Y) = \sum_i \sum_j p_{ij} \ln \frac{p_{ij}}{p_{i\cdot} p_{\cdot j}}$$

用频率代替概率，得到

$$\chi_L^2 = 2\sum_i \sum_j f_{ij} \ln \frac{n f_{ij}}{f_{i\cdot} f_{\cdot j}} \text{①} \qquad (10.33)$$

式中，f_{ij} 是第 ij 格的频数，$f_{i\cdot}$ 是第 i 行的频数和，$f_{\cdot j}$ 是第 j 列的频数和，即为边缘次数，n 是样本总数。χ^2 似然比统计量遵从自由度 $df = (r-1)(c-1)$ 的 χ^2 分布。在附表 I 中，可以根据 df 查到变量相互独立的假设成立时 χ^2 为某值的概率，以判定模型的拟合效果。若概率越大，表明模式的拟合效果越好。在多元线性回归中，采用拟合优度 R^2 考察模型的拟合效果，R^2 越大越好。对数线性模型采用 χ_L^2 检验模型的拟合效果，χ_L^2 越小越好。

3. 三个变量的 χ^2 似然比

检验三个变量所建立的对数线性模型拟合效果，仍然可以利用(10.32)式的 χ^2 似然比统计量。但由于变量间关系不同，相应的概率估计不同，即 p_{ijk} 的计算不同，因而联合熵的计算也有差异。根据表 10—7、表 10—8 的变量关系和模型类型分别讨论如下。

Ⅰ.(XYZ)。三个变量相互不独立，建立饱和模型。模型对数据完全拟合，$\chi_L^2 = 0$。若利用列联表频数计算，可直接代入(10.32)式求得联合熵。

Ⅱ.(X,YZ)。这时，X 与(Y,Z)独立。有 $P_{ijk} = p_{i\cdot\cdot} p_{\cdot jk}$，所以联合熵的估计为

$$\hat{H} = \hat{H}_x + \hat{H}_{yz}$$
$$= -\sum_i p_{i\cdot\cdot} \ln p_{i\cdot\cdot} - \sum_j \sum_k p_{\cdot jk} \ln p_{\cdot jk}$$

① 设 $G = H_{独立} - H$，则 $2nG = -2\ln\Lambda$

$$= -\sum_i \frac{f_{i\cdot\cdot}}{n} \ln \frac{f_{i\cdot\cdot}}{n} - \sum_j \sum_k \frac{f_{\cdot jk}}{n} \ln \frac{f_{\cdot jk}}{n} \quad (10.34)$$

若是(Y,XZ)或(Z,XY)的情况，联合熵的估计与其类似，只是注意概率间的关系变化，因而观测频数f_{ijk}的下标也发生相应的变化。

Ⅲ.(XY,XZ)。这时是条件独立的情况，用条件概率较为方便。由于

$$\begin{aligned} p_{ijk} &= P(X=i)\cdot P=(Y=j|X=i)\cdot P((Z=k|X=i) \\ &= p_{i\cdot\cdot}\frac{p_{ij\cdot}}{p_{i\cdot\cdot}}\cdot\frac{p_{i\cdot k}}{p_{i\cdot\cdot}} \\ &= \frac{p_{ij\cdot}\,p_{i\cdot k}}{p_{i\cdot\cdot}} \end{aligned}$$

所以，联合熵的估计为

$$\begin{aligned} \hat{H} &= -\sum_i\sum_j\sum_k \frac{p_{ij\cdot}\,p_{i\cdot k}}{p_{i\cdot\cdot}} \ln \frac{P_{ij\cdot}\,p_{i\cdot k}}{p_{i\cdot\cdot}} \\ &= -\sum_i\sum_j\sum_k \frac{f_{ij\cdot}\,f_{i\cdot k}}{nf_{i\cdot\cdot}} \ln \frac{f_{ij\cdot}\,f_{i\cdot k}}{n\,f_{i\cdot\cdot}} \\ &= -\sum_i \frac{f_{i\cdot\cdot}}{n} \ln \frac{f_{i\cdot\cdot}}{n} - \sum_i\sum_j \frac{f_{ij\cdot}}{f_{i\cdot\cdot}} \ln \frac{f_{ij\cdot}}{f_{i\cdot\cdot}} \\ &\quad - \sum_i\sum_k \frac{f_{i\cdot k}}{f_{i\cdot\cdot}} \ln \frac{f_{i\cdot k}}{f_{i\cdot\cdot}} \end{aligned} \quad (10.35)$$

这一类的其余模型可以采用类似方法求得联合熵。

Ⅳ.(X,Y,Z)。这是三个变量相互独立的情况。有$p_{ijk} = p_{i\cdot\cdot}\,p_{\cdot j\cdot}\,p_{\cdot\cdot k}$，联合熵的估计为

$$\begin{aligned} \hat{H} &= \hat{H}_X + \hat{H}_Y + \hat{H}_Z \\ &= -\sum_i \frac{f_{i\cdot\cdot}}{n} \ln \frac{f_{i\cdot\cdot}}{n} - \sum_j \frac{f_{\cdot j\cdot}}{n} \ln \frac{f_{\cdot j\cdot}}{n} \\ &\quad - \sum_k \frac{f_{\cdot\cdot k}}{n} \ln \frac{f_{\cdot\cdot k}}{n} \end{aligned} \quad (10.36)$$

根据似然比的含义及两个变量的χ^2似然比计算可知，对于三个变量的χ^2似然比，也可以利用变量相互独立时的联合熵与欲建立的模型所对应的联合熵之差得到(熵的变化与似然比的变化是等价的)。这一计算比较麻烦，好在电子计算机的运用提供了方便，

特别是一些统计应用软件如 SPSS、SAS 等，都有对数线性分析的功能，能够直接运用。

【例 10.7】 性别、任现职与转业意愿的关系分析

在某地区对某些人进行调查，结果如表 10—20。分析研究三个变量间关系。

分析：这三个变量都是定类尺度测量的变量，属于定性变量。三个变量间关系的分析可以采用对数线性分析方法。就三个变量可能的所有关系分别建立对数线性模型，计算每个模型的 χ_L^2，根据自由度，在附表 I 中查找 H_0：三个变量相互独立成立时，似然比统计量（$-2\ln\Lambda$ 即 χ_L^2）为某值的概率。概率值越大，表明 χ_L^2 越显著，独立模型应拒绝，而该模型拟合效果越好。χ_L^2 自由度由（10.37）式确定。

表 10—20 **性别、现职及转业意愿调查结果**

性别	现职	转业意愿(Z)	
(X)	(Y)	有	无
女	非农业劳动	19	132
	农业劳动	0	9
男	非农业劳动	11	52
	农业劳动	6	97

$$\mathrm{df} = \text{表上单元数目} - \text{模型拟合中需估计的参数数目} \quad (10.37)$$

对表 10—20 的数据进行各种拟合，结果如表 10—21。其中自由度按（10.37）式计算。如对于（X,Y,Z）由表 10—8 可知，模型需估计的参数为 λ、λ_i^1、λ_j^2、λ_k^3，共计 4 个，表中单元数目为 8，则自由度 $\mathrm{df}=8-4=4$。其它模型 χ^2 似然比统计量的自由度可以类推。

表 10—21 的结果表明，模型（YZ,XY）的拟合效果最好，其 P 值最大，说明在 H_0 成立时 χ_L^2 大于等于表中数值的概率很大。其余模型拟合均不理想。由模型拟合效果也可以看到变量间的关系。表 10—21 中，凡是忽略 X 与 Y 间关系的模型，其 χ_L^2 的 P 值均为 0，拟合效果极差，表明性别与现职的相关性不容忽视。同时，可以看出，

X 与 Z 之间没有直接的相关性，凡包含(XZ)的模型，其 χ^2_L 的 P 也很小，甚至为 0，表明性别和转业意愿之间没有相关性。模型(YZ，XY) 拟合得好，表明三个变量间是条件独立性关系，在 Y 的每个水平上，X 与 Z 不相关，Y 与 X 和 Z 分别相关，也就是说，性别与现职相关，现职与转业意愿相关。

表 10—21　　　　各模型的拟合效果[①]

模型	χ^2_L	df	P 值
(X,Y,Z)	137.93	4	0.000
(X,YZ)	131.68	3	0.000
(Y,XZ)	137.71	3	0.000
(Z,XY)	8.13	3	0.043
(XZ,YZ)	131.46	2	0.000
(XZ,XY)	7.91	2	0.019
(YZ,XY)	1.88	2	0.390
(XYZ)	0	0	

(二) 残差分析

饱和模型对联列表数据完全拟合，因而不存在残差，即残差为 0，但对于其它模型，就不能保证每单元的观测值与期望值完全相等，因而有必要进行残差分析。

残差用以表明模型对数据的拟合程度。残差越小，表明模型的拟合程度越高。

1. *原始残差*

原始残差即列联表每个单元的残差，其计算公式为

$$残差 = 该单元观测值 - 期望值 \qquad (10.38)$$

对于二维表来说，残差 RESID 为

$$\text{RESID} = f_{ij} - e_{ij}$$

对于三维表来说有

$$\text{RESID} = f_{ijk} - e_{ijk}$$

① 表中结果由社会统计应用软件包(SPSS)计算得到，以后的计算均采用该软件

由于实际观测值可能大于或小于期望值，因而残差可能为正，可能为负。无论正负，残差越接近于0，表明模型的拟合程度越高。

2. 标准化残差

为使用时更方便，人们对原始残差给予各种方式的校正，其中 Pearson 标准化残差应用较为广泛。其计算公式为

$$标准化残差 = \frac{观测值 - 期望值}{\sqrt{期望值}} \qquad (10.39)$$

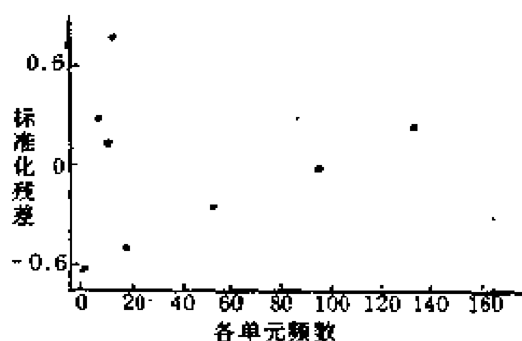


图 10.1 各单元标准化残差与观测值对应图

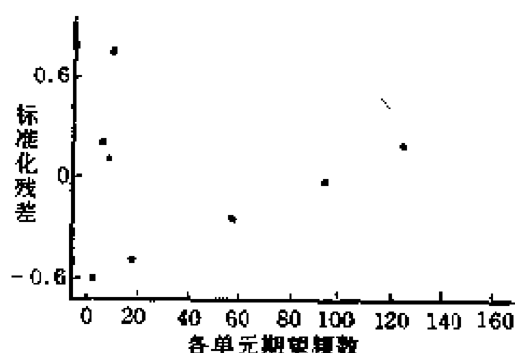


图 10.2 各单元标准化残差与期望值对应图

若模型选择正确，则标准化残差接近正态分布。一般标准化残差在 ± 1.96 范围内表明残差不大，模型选择合适，拟合效果为好。表 10—22 是例 10.7 中，模型(XY,YZ) 各单元的原始残差和标准化残差。由表可知，标准化残差全部落在 ± 0.80 内，表明残差很小，模型的拟合效果确实很好。如果标准化残差近似正态分布，那么各单元标准化残差与观测值、期望值在直角坐标图上形成直线排列。图 10.1 是表 10—22 各单元观测频数与标准化残差的对应图。图 10.2 是各单元期望频数与标准化残差的对应图。由图可以看出，

它们的点分布差异不大，表明模型适合这些数据。

表 10—22　　模型(XY,YZ)的残差

单元频数	单元期望频数	原始残差	标准化残差
19	21.17	－2.17	－0.47
132	129.83	2.17	0.19
0	0.48	－.48	－0.69
9	8.52	0.48	0.17
11	8.83	2.17	0.73
52	54.17	－2.17	－0.29
6	5.52	0.48	0.21
97	97.48	－0.48	－0.05

二、模型的选择

对于一批数据来说，用什么样的对数线性模型拟合最好，即用哪类模型能最恰当地反映变量之间的关系，需要对模型进行选择。模型选择实际上就是寻找、识别最合适的模型。对数线性模型选择的方法主要有：系数选择法、χ^2 似然比法、自后淘汰法等。

(一) 系数选择法

回归分析法中判断所选自变量是否合适，或说是否建立的为最优回归方程，常常采用回归系数的 t 检验法。对数线性分析与其类似，模型建立后，各个效应的估计值是否显著也应进行检验。通过效应系数的显著性可以判断该效应项包含在模型中是否合理。对数线性模型系数的检验通常用 Z 检验法。Z 值为

$$Z=\frac{\text{系数估计值}}{\text{系数标准差}}$$

若 $|Z|>1.96$，表明该系数显著不为 0，该效应项应在模型中保留；若 $|Z|<1.96$，则系数与 0 无显著差异，该效应项需删除。

【例 10.8】　采用系数选择法为表 10－20 的数据选择合适的对数线性模型

分析：采用系数选择法确定合适的模型，首先应建立饱和模型，然后对每个效应系数进行显著性检验。利用表 10—20 的数据

建立饱和模型，得到各个主效应以及交互效应的估计值、Z 值，如表 10—23，由于总平均效应在各类模型中必不可少，因而检验时可不考虑。

从表 10—23 的结果可知，交互效应项 λ_{11}^{13} 的系数估计值的绝对值很小，相应的 $|Z| = 0.4064 < 1.96$，表明该效应项不显著，其为 0 的可能性极大，因而将其包含在模型中不合适。在谱系模型中，低阶效应项为 0，相应的高阶效应项必然为 0，因此，$\lambda_{111}^{123} = 0$。由表 10—23 也可知，二次交互作用效应项 λ_{111}^{123} 的系数估计值绝对值很小，$|Z| = 0.1034 < 1.96$，该效应项确实也应删去。对于交互效应项 λ_{11}^{23} 是否要删除，应取慎重态度。从 Z 值看，$|Z| = 1.4215 < 1.96$，似乎并不显著，但系数估计值为 0.2772 并不很小，这时应结合其他方法考察。

表 10—23 各主效应、交互效应估计值、Z 值

效应	系数估计值	Z 值
λ_1^1	−0.4300	−2.2047
λ_1^2	0.7813	4.0062
λ_1^3	−1.1359	−5.8244
λ_{11}^{12}	0.7934	4.0683
λ_{11}^{13}	−0.0793	−0.4064
λ_{11}^{23}	0.2772	1.4215
λ_{111}^{123}	0.0202	−0.1034

(二)χ^2 似然比法

χ^2 似然比法亦称模型分块选择法。它是利用 χ^2 似然比的性质来选择适宜的对数线性模型的方法。在回归分析中，当模型增加一个新的变量时，若拟合优度 R^2 增大，表明该变量对模型有贡献。这说明 R^2 的增加量刻划了新变量所提供的附加信息。在对数线性分析中，χ^2 似然比的减少量可以刻划变量对模型贡献的大小。若模型引入一个效应项，使 χ^2 似然比减少，则该效应项对模型有贡献；反

之，引入的效应项使 χ^2 似然比增大，则该效应项不应引入模型。对于对数线性模型来说，χ^2 似然比越小，表明模型越好。含有不同效应项的两个模型 χ^2 似然比之差值，恰恰反映了两个模型中不同效应项的贡献。

采用 χ^2 似然比法选择模型，一般是设计几种除饱和模型以外的模型，分别计算 χ^2 似然比，根据 χ^2 似然比的变动作为模型选择的依据。设计模型时，可以从独立模型，即只有主效应的模型开始，依次引入各个交互作用效应项；也可以从包含效应项最多的模型开始，依次减少各个效应项。

【例 10.9】 采用 χ^2 似然比法为表 10 — 20 的数据选择合适的对数线性模型

分析：从独立模型 (X,Y,Z) 开始，依次设计几个模型，如表 10—21 所列。独立模型 (X,Y,Z) 的 $\chi_L^2 = 137.93$，相应的 $P = 0.0 < 0.05$，否定没有交互效应项的假设，引入交互效应项 (YZ)，模型 (X,YZ) 的 $\chi_L^2 = 131.68$，其与独立模型 χ^2 似然比的差值为 $137.93 - 131.68 = 6.25$，引入 YZ 交互效应使 χ_L^2 减少，表明该效应项对模型有贡献。但模型 (X,YZ) 的 χ_L^2 显著性水平 $P = 0.0 < 0.05$，表明模型仅有 YZ 交互效应项的假设应予以否定。模型中引入 XZ 项，$\chi_L^2 = 137.71$，比模型引入 YZ 项的 χ_L^2 增大，比独立模型的 χ_L^2 只减少 0.22，表明交互效应项 XZ 对模型几乎没有贡献，(Y,XZ) 模型应予以否定。模型若引入 XY 交互效应项，$\chi_L^2 = 8.13$，比独立模型的 χ_L^2 似然比大大减少，表明交互效应项 XY 对模型有贡献，但模型 (Z,XY) χ_L^2 的显著性水平 $P = 0.043 < 0.05$，仍需否定。依照上面的方法，对后三个模型进行判定，最后一个模型 (XZ,XY) 的 χ_L^2 最小，仅为 1.88，与独立模型的 $\chi_L^2 = 137.93$ 相比，减少 $137.93 - 1.88 = 136.05$，差值很大，表明交互效应项 YZ、XY 对模型贡献很大，χ_L^2 的显著水平 $P = 0.39 > 0.05$，表明模型包含这两个交互效应项的假设不能被否定。利用 χ^2 似然比法选择的最优对数线性模型为 (YZ,XY)，即 $\lambda + \lambda_i^1 - \lambda_j^2 + \lambda_k^3 + \lambda_{ij}^{12} + \lambda_{jk}^{23}$。

（三）自后淘汰法

自后淘汰法亦称向后删除法。它是利用变量选取模型的方法。开始将所有效应包含在模型之中，可以建立饱和模型，也可以是任何一个谱系模型。在谱系模型中，变量的高阶效应存在，其低阶效应必存在，如模型中若存在三个变量的交互效应项 λ_{ijk}^{123}，则必然有 λ_{ij}^{12}、λ_{ik}^{13}、λ_{jk}^{23} 以及 λ_i^1、λ_j^2、λ_k^3。因此，描述谱系模型只需用最高阶效应项表示即可，这被称作“模型的生成类”。如表 10 － 8 中，(XYZ) 表示模型中含有 λ_{ijk}^{123} 以及所有较低阶的其他效应项，而(Z)、(XY) 表示模型 $\lambda + \lambda_i^1 + \lambda_j^2 + \lambda_k^3 + \lambda_{ij}^{12}$。因为$(XY)$ 是第 1 个变量 X 与第 2 个变量 Y 的最高阶效应项，而 Z 只以主效应存在。自后淘汰法是从模型的最高阶效应开始，依次逐个淘汰使 χ_L^2 变化最小的效应，换句话说，是通过 χ_L^2 检验对应于生成类的效应是否为 0，逐个删除那些在给定的显著性水平上（一般选为 0.05）不能否定为 0 的效应项；若有多个项需删除，则删除其中最大显著水平对应的项，直至得到最好生成类。

【例10.10】 采用自后淘汰法为表 10 － 20 的数据选择最适宜的对数线性模型

分析：首先建立饱和模型(XYZ)[①]，得到 $\chi_L^2 = 0$，若删除最高阶效应项 λ_{ijk}^{123}，也就是生成类(XYZ)，则 $\chi_L^2 = 0.701$，自由度 df ＝ 1，概率 $P = 0.4024$，这表明生存类(XYZ) 为 0 的概率是 $0.4024 > 0.05$，不能否定该项效应为 0，其对模型的影响不显著，故而应删除，得到第一步的最好生成类(XY, XZ, YZ)。这是最高阶效应为三个两两交互效应项的模型。由于有三个生成类(XY)、(XZ)、(YZ)，故应分别考察，计算结果如表 10 － 24。由表可以看出，生成类(XY) 效应项为 0 的概率 $P = 0.2772 > 0.05$，不能否定为 0，故应删除。若将其删除，重新建模 $\chi_L^2 = 1.8819$，$P = 0.39 > 0.05$，表明该效应项对模型的影响不显著，可以删除，得到第二步的最好生

① SPSS 软件中有自后淘汰法功能，本例采用该软件完成。

表 10—24　　第一步生成类的检验值

检验项	自由度(df)	χ_L^2	P
XY	1	130.757	0.0000
XZ	1	1.181	0.2772
YZ	1	7.209	0.0073

成类(XY,YZ)。对第二步生成类的两项(XY)、(YZ)分别进行检验，计算结果如表 10－25。由表可知，两个生成类项为 0 的概率均小于 5%，否定该项为 0 的假设。这样，用自后淘汰法得到的最好模型为(XY,YZ)。

例 10.8、例 10.9、例 10.10 分别采用三种方法为表 10－20 的数据建立对数线性模型，结果都选择了模型(XY,YZ)，即

$$\ln e_{ijk} = \lambda + \lambda_i^1 + \lambda_j^2 + \lambda_k^3 + \lambda_{ij}^{12} + \lambda_{jk}^{23}$$

这表明性别和现职，现职和转业意愿存在相关，而性别和转业意愿关系不大。

表 10—25　　第二步生成类的检验值

检验项	df	χ_L^2	P
XY	1	129.798	0.0000
YZ	1	6.250	0.0124

例 10.7 对模型(XY,YZ)进行了拟合程度的检验，$\chi_L^2 = 1.8819$，$P = 0.390 > 0.05$，表明模型对数据总的拟合程度很好。表 10—22 显示，模型得到的各单元期望值与实际观测值之间误差的最大绝对值为 2.17，最大标准化残差绝对值为 0.73，没有超过 1.96，说明每个单元模型拟合数据程度很高。模型通过了拟合程度的检验，表明该模型确实是对数据拟合最好的模型。

附表 I：χ^2 分布表

	右尾概率													
df	.99	.98	.95	.90	.80	.70	.50	.30	.20	.10	.05	.02	.01	.001
1	.00016	.00063	.0039	.016	.064	.15	.46	1.07	1.64	2.71	3.84	5.41	6.64	10.83
2	.02	.04	.10	.21	.45	.71	1.39	2.41	3.22	4.60	5.99	7.82	9.21	13.82
3	.12	.18	.35	.58	1.00	1.42	2.37	3.66	4.64	6.25	7.82	9.84	11.34	16.27
4	.30	.43	.71	1.06	1.65	2.20	3.36	4.88	5.99	7.78	9.49	11.67	13.28	18.46
5	.55	.75	1.14	1.61	2.34	3.00	4.35	6.06	7.29	9.24	11.07	13.39	15.09	20.52
6	.87	1.13	1.64	2.20	3.07	3.83	5.35	7.23	8.56	10.64	12.59	15.03	16.81	22.46
7	1.24	1.56	2.17	2.83	3.82	4.67	6.35	8.38	9.80	12.02	14.07	16.62	18.48	24.32
8	1.65	2.03	2.73	3.49	4.59	5.53	7.34	9.52	11.03	13.36	15.51	18.17	20.09	26.12
9	2.09	2.53	3.32	4.17	5.38	6.39	8.34	10.66	12.24	14.68	16.92	19.68	21.67	27.88
10	2.56	3.06	3.94	4.86	6.18	7.27	9.34	11.78	13.44	15.99	18.31	21.16	23.21	29.59
11	3.05	3.61	4.58	5.58	6.99	8.15	10.34	12.90	14.63	17.28	19.68	22.62	24.72	31.26
12	3.57	4.18	5.23	6.30	7.81	9.03	11.34	14.01	15.81	18.55	21.03	24.05	26.22	32.91
13	4.11	4.76	5.89	7.04	8.63	9.93	12.34	15.12	16.98	19.81	22.36	25.47	27.69	34.53
14	4.66	5.37	6.57	7.79	9.47	10.82	13.34	16.22	18.15	21.06	23.68	26.87	29.14	36.12
15	5.23	5.89	7.26	8.55	10.31	11.72	14.34	17.32	19.31	22.31	25.00	28.26	30.58	37.70

(续附表Ⅰ)

	右尾概率													
df	.99	.98	.95	.90	.80	.70	.50	.30	.20	.10	.05	.02	.01	.001
16	5.81	6.61	7.96	9.31	11.15	12.62	15.34	18.42	20.46	23.54	26.30	29.63	32.00	39.29
17	6.41	7.26	8.67	10.08	12.00	13.53	16.34	19.51	21.62	24.77	27.59	31.00	33.41	40.75
18	7.02	7.91	9.39	10.86	12.86	14.44	17.34	20.60	22.76	25.99	28.87	32.35	34.80	42.31
19	7.63	8.57	10.12	11.65	13.72	15.35	18.34	21.69	23.90	27.20	30.14	33.69	36.19	43.82
20	8.26	9.24	10.85	12.44	14.58	16.27	19.34	22.78	25.04	28.41	31.41	35.02	37.57	45.32
21	8.90	9.92	11.59	13.24	15.44	17.18	20.34	23.86	26.17	29.62	32.67	36.34	38.93	46.80
22	9.54	10.60	12.34	14.04	16.31	18.10	21.34	24.94	27.30	30.81	33.92	37.66	40.29	48.27
23	10.20	11.29	13.09	14.58	17.10	19.02	22.34	26.02	28.43	32.01	35.17	38.97	41.64	49.73
24	10.86	11.99	13.85	15.66	18.06	19.94	23.34	27.10	29.55	33.20	36.42	40.27	42.98	51.18
25	11.52	12.70	14.61	16.47	18.94	20.87	24.34	28.17	30.68	34.38	37.65	41.57	44.31	52.62
26	12.20	13.41	15.38	17.29	19.82	21.79	25.34	29.25	31.80	35.56	38.88	42.86	45.64	54.05
27	12.88	14.12	16.15	18.11	20.70	22.72	26.34	30.32	32.91	36.74	40.11	44.14	46.96	55.48
28	13.56	14.86	16.93	18.94	21.59	23.65	27.34	31.39	34.03	37.92	41.34	45.42	48.28	56.89
29	14.26	15.57	17.71	19.77	22.48	24.58	28.34	32.46	35.14	39.09	42.56	46.69	49.59	58.30
30	14.95	16.31	18.49	20.60	23.36	25.51	29.34	33.53	36.25	40.26	43.77	47.96	50.89	59.70

如果 $df>30$，那么可以采用下面公式得到 Z，在附表Ⅳ中当找到近似的右尾或左尾概率：$Z=\sqrt{2Q}-\sqrt{2(df)-1}$

附表Ⅱ：二项分布表

n	x	θ .05	.10	.15	.20	.25	.30	.35	.40	.45
1	0	.9500	.9000	.8500	.8000	.7500	.7000	.6500	.6000	.5500
	1	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2	0	.9025	.8100	.7225	.6400	.5625	.4900	.4225	.3600	.3025
	1	.9975	.9900	.9775	.9600	.9375	.9100	.8775	.8400	.7975
	2	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
3	0	.8574	.7290	.6141	.5120	.4219	.3430	.2746	.2160	.1664
	1	.9928	.9720	.9392	.8960	.8438	.7840	.7182	.6480	.5748
	2	.9999	.9990	.9966	.9920	.9844	.9730	.9571	.9360	.9089
	3	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
4	0	.8145	.6561	.5220	.4096	.3164	.2401	.1785	.1296	.0915
	1	.9860	.9477	.8905	.8192	.7383	.6517	.5630	.4752	.3910
	2	.9995	.9963	.9880	.9728	.9492	.9163	.8735	.8208	.7585
	3	1.0000	.9999	.9995	.9984	.9961	.9919	.9850	.9744	.9590
	4	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
5	0	.7738	.5905	.4437	.3277	.2373	.1681	.1160	.0778	.0503
	1	.9774	.9185	.8352	.7373	.6328	.5282	.4284	.3370	.2562
	2	.9988	.9914	.9734	.9421	.8965	.8369	.7648	.6826	.5931
	3	1.0000	.9995	.9978	.9933	.9844	.9692	.9460	.9130	.8688
	4	1.0000	1.0000	.9999	.9997	.9990	.9976	.9947	.9898	.9815
	5	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
6	0	.7351	.5314	.3771	.2621	.1780	.1176	.0754	.0467	.0277
	1	.9672	.8857	.7765	.6554	.5339	.4202	.3191	.2333	.1636
	2	.9978	.9842	.9527	.9011	.8306	.7443	.6471	.5443	.4415
	3	.9999	.9987	.9941	.9830	.9624	.9295	.8826	.8208	.7447
	4	1.0000	.9999	.9996	.9984	.9954	.9891	.9777	.9590	.9308
	5	1.0000	1.0000	1.0000	.9999	.9998	.9993	.9982	.9959	.9917
	6	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
7	0	.6983	.4783	.3206	.2097	.1335	.0824	.0490	.0280	.0152
	1	.9556	.8503	.7166	.5767	.4449	.3294	.2338	.1586	.1024
	2	.9962	.9743	.9262	.8520	.7564	.6471	.5323	.4199	.3164
	3	.9998	.9973	.9879	.9667	.9294	.8740	.8002	.7102	.6083
	4	1.0000	.9998	.9988	.9953	.9871	.9712	.9444	.9037	.8471
	5	1.0000	1.0000	.9999	.9996	.9987	.9962	.9910	.9812	.9643
	6	1.0000	1.0000	1.0000	1.0000	.9999	.9998	.9994	.9984	.9963
	7	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

（续附表 I ）

n	x					θ					
		.50	.55	.60	.65	.70	.75	.80	85	.90	.95
1	0	.5000	.4500	.4000	.3500	.3000	.2500	.2000	.1500	.1000	.0500
	1	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2	0	.2500	.2025	.1600	.1225	.0900	.0625	.0400	.0225	.0100	.0025
	1	.7500	.6975	.6400	.5775	.5100	.4375	.3600	.2775	.1900	0975
	2	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
3	0	.1250	.0911	.0640	.0429	.0270	.0156	.0080	.0034	.0010	.0001
	1	.5000	.4252	.3520	.2818	.2160	.1562	.1040	.0608	.0280	.0072
	2	.8750	.8336	.7840	.7254	.6570	.5781	.4880	.3859	.2710	.1426
	3	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
4	0	.0625	.0410	.0256	.0150	.0081	.0039	.0016	.0005	.0001	.0000
	1	.3125	.2415	.1792	.1265	.0837	.0508	.0272	.0120	.0037	.0005
	2	.6875	.6090	.5248	.4370	.3483	.2617	.1808	.1095	.0523	.0140
	3	.9375	.9085	.8704	.8215	.7599	.6836	.5904	.4780	.3439	.1855
	4	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
5	0	.0312	.0185	.0102	.0053	.0024	.0010	.0003	.0001	.0000	.0000
	1	.1857	1312	.0870	.0540	.0308	.0156	.0067	.0022	.0005	.0000
	2	.5000	.4069	.3174	.2352	.1631	.1035	.0579	.0266	.0086	.0012
	3	.8125	.7438	.6630	.5716	.4718	.3672	.2627	.1648	.0815	.0226
	4	9688	.9497	.9222	.8840	.8319	.7627	.6723	.5563	.4095	.2262
	5	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
6	0	.0156	.0083	.0041	.0018	.0007	.0002	.0001	.0000	.0000	.0000
	1	.1094	.0692	.0410	.0223	.0109	.0046	.0016	.0004	.0001	.0000
	2	.3438	.2553	.1792	.1174	.0705	.0376	.0170	.0059	.0013	.0001
	3	.6562	.5585	.4557	.3529	.2557	.1694	.0989	.0473	.0158	.0022
	4	.8906	.8364	.7667	.6809	.5798	.4661	.3446	.2235	.1143	.0328
	5	.9844	.9723	.9533	.9246	.8824	.8220	.7379	.6229	.4686	.2649
	6	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
7	0	.0078	.0037	.0016	.0006	.0002	.0001	.0000	.0000	.0000	.0000
	1	.0625	.0357	.0188	.0090	.0038	.0013	.0004	.0001	.0000	.0000
	2	.2266	.1529	.0963	.0556	.0288	.0129	.0047	.0012	.0002	.0000
	3	.5000	.3917	.2898	.1998	.1260	.0706	.0333	.0121	.0027	.0002
	4	.7734	.6836	.5801	.4677	.3529	.2436	.1480	.0738	.0257	.0038
	5	.9375	.8976	.8414	.7662	.6706	.5551	.4233	.2834	.1497	.0444
	6	.9922	.9848	.9720	.9510	.9176	.8665	.7903	.6794	.5217	.3017
	7	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

(续附表 I)

n	x	θ .05	.10	.15	.20	.25	.30	.35	.40	.45
12	0	.5404	.2824	.1422	.0687	.0317	.0138	.0057	.0022	.0008
	1	.8816	.6590	.4435	.2749	.1584	.0850	.0424	.0196	.0083
	2	.9804	.8891	.7358	.5583	.3907	.2528	.1513	.0834	.0421
	3	.9978	.9744	.9078	.7946	.6488	.4925	.3467	.2253	.1345
	4	.9998	.9957	.9761	.9274	.8424	.7237	.5833	.4382	.3044
	5	1.0000	.9995	.9954	.9806	.9456	.8822	.7873	.6652	.5269
	6	1.0000	.9999	.9993	.9961	.9857	.9614	.9154	.8418	.7393
	7	1.0000	1.0000	.9999	.9994	.9972	.9905	.9745	.9427	.8883
	8	1.0000	1.0000	1.0000	.9999	.9996	.9983	.9944	.9847	.9644
	9	1.0000	1.0000	1.0000	1.0000	1.0000	.9998	.9992	.9972	.9921
	10	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9997	.9989
	11	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999
	12	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
13	0	.5133	.2542	.1209	.0550	.0238	.0097	.0037	.0013	.0004
	1	.8646	.6213	.3983	.2336	.1267	.0637	.0296	.0126	.0049
	2	.9755	.8661	.7296	.5017	.3326	.2025	.1132	.0579	.0269
	3	.9969	.9658	.9033	.7473	.5843	.4206	.2783	.1686	.0929
	4	.9997	.9935	.9740	.9009	.7940	.6543	.5005	.3530	.2279
	5	1.0000	.9991	.9947	.9700	.9198	.8346	.7159	.5744	.4268
	6	1.0000	.9999	.9987	.9930	.9757	.9376	.8705	.7712	.6437
	7	1.0000	1.0000	.9998	.9988	.9944	.9818	.9538	.9023	.8212
	8	1.0000	1.0000	1.0000	.9998	.9990	.9960	.9874	.9679	.9302
	9	1.0000	1.0000	1.0000	1.0000	.9999	.9993	.9975	.9922	.9797
	10	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9997	.9987	.9959
	11	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9995
	12	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	13	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
14	0	.4877	.2288	.1028	.0440	.0178	.0068	.0024	.0008	.0002
	1	.8470	.5846	.3567	.1979	.1010	.0475	.0205	.0081	.0029
	2	.9699	.8416	.6479	.4481	.2811	.1608	.0839	.0398	.0170
	3	.9958	.9559	.8535	.6982	.5213	.3552	.2205	.1243	.0632
	4	.9996	.9908	.9533	.8702	.7415	.5842	.4227	.2793	.1672
	5	1.0000	.9985	.9885	.9561	.8883	.7805	.6405	.4859	.3373
	6	1.0000	.9998	.9978	.9884	.9617	.9067	.8164	.6925	.5461
	7	1.0000	1.0000	.9997	.9976	.9897	.9685	.9247	.8499	.7414
	8	1.0000	1.0000	1.0000	.9996	.9978	.9917	.9757	.9417	.8811
	9	1.0000	1.0000	1.0000	1.0000	.9997	.9983	.9940	.9825	.9574
	10	1.0000	1.0000	1.0000	1.0000	1.0000	.9998	.9989	.9961	.9886
	11	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9994	.9978
	12	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9997
	13	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

（续附表 I）

n	x	.50	.55	.60	.65	θ .70	.75	.80	85	.90	.95
12	0	.0002	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	1	.0032	.0011	.0003	.0001	.0000	.0000	.0000	.0000	.0000	.0000
	2	.0193	.0079	.0028	.0008	.0002	.0000	.0000	.0000	.0000	.0000
	3	.0730	.0356	.0153	.0056	.0017	.0004	.0001	.0000	.0000	.0000
	4	.1938	.1117	.0573	.0255	.0095	.0028	.0006	.0001	.0000	.0000
	5	.3872	.2607	.1582	.0846	.0386	.0143	.0039	.0007	.0001	.0000
	6	.6128	.4731	.3348	.2127	.1178	.0544	.0194	.0046	.0005	.0000
	7	.8062	.6956	.5618	.4167	.2763	.1576	.0726	.0239	.0043	.0002
	8	.9270	.8655	.7747	.6533	.5075	.3512	.2054	.0922	.0256	.0022
	9	.9807	.9579	.9166	.8487	.7472	.6093	.4417	.2642	.1109	.0196
	10	.9968	.9917	.9804	.9576	.9150	.8416	.7251	.5565	.3410	.1184
	11	.9998	.9992	.9978	.9943	.9862	.9683	.9313	.8578	.7176	.4596
	12	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
13	0	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	1	.0017	.0005	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	2	.0112	.0041	.0013	.0003	.0001	.0000	.0000	.0000	.0000	.0000
	3	.0461	.0203	.0078	.0025	.0007	.0001	.0000	.0000	.0000	.0000
	4	.1334	.0698	.0321	.0126	.0040	.0010	.0002	.0000	.0000	.0000
	5	.2905	.1788	.0977	.0462	.0182	.0056	.0012	.0002	.0000	.0000
	6	.5000	.3563	.2288	.1295	.0624	.0243	.0070	.0013	.0001	.0000
	7	.7095	.5732	.4256	.2841	.1654	.0802	.0300	.0053	.0009	.0000
	8	.8666	.7721	.6470	.4995	.3457	.2060	.0991	.0260	.0065	.0003
	9	.9539	.9071	.8314	.7217	.5794	.4157	.2527	.0967	.0342	.0031
	10	.9888	.9731	.9421	.8868	.7975	.6674	.4983	.2704	.1339	.0245
	11	.9983	.9951	.9874	.9704	.9363	.8733	.7664	.6017	.3787	.1354
	12	.9999	.9996	.9987	.9963	.9903	.9762	.9450	.8791	.7458	.4867
	13	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
14	0	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	1	.0009	.0003	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	2	.0065	.0022	.0006	.0001	.0000	.0000	.0000	.0000	.0000	.0000
	3	.0287	.0114	.0039	.0011	.0002	.0000	.0000	.0000	.0000	.0000
	4	.0898	.0462	.0175	.0060	.0017	.0003	.0000	.0000	.0000	.0000
	5	.2120	.1189	.0583	.0243	.0083	.0022	.0004	.0000	.0000	.0000
	6	.3953	.2586	.1501	.0753	.0315	.0103	.0024	.0003	.0000	.0000
	7	.6047	.4539	.3075	.1836	.0933	.0383	.0116	.0022	.0002	.0000
	8	.7880	.6627	.5141	.3595	.2195	.1117	.0439	.0115	.0015	.0000
	9	.9102	.8328	.7207	.5773	.4158	.2585	.1298	.0467	.0092	.0004
	10	.9713	.9368	.8757	.7795	.6448	.4787	.3018	.1465	.0441	.0042
	11	.9935	.9830	.9602	.9161	.8392	.7189	.5519	.3521	.1584	.0301
	12	.9991	.9971	.9919	.9795	.9525	.8990	.8021	.6433	.4154	.1530
	13	.9999	.9998	.9992	.9976	.9932	.9822	.9560	.8972	.7712	.5123
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

(续附表Ⅰ)

n	x	.05	.10	.15	.20	θ .25	.30	.35	.40	.45
15	0	.4633	.2059	.0874	.0352	.0134	.0047	.0016	.0005	.0001
	1	.8290	.5490	.3186	.1671	.0802	.0353	.0142	.0052	.0017
	2	.9638	.8159	.6042	.3980	.2361	.1268	.0617	.0271	.0107
	3	.9945	.9444	.8227	.6482	.4613	.2969	.1727	.0905	.0424
	4	.9994	.9873	.9383	.8358	.6865	.5155	.3519	.2173	.1204
	5	.9999	.9978	.9832	.9389	.8516	.7216	.5643	.4032	.2608
	6	1.0000	.9997	.9964	.9819	.9434	.8689	.7548	.6098	.4522
	7	1.0000	1.0000	.9994	.9958	.9827	.9500	.8868	.7869	.6535
	8	1.0000	1.0000	.9999	.9992	.9958	.9848	.9578	.9050	.8182
	9	1.0000	1.0000	1.0000	.9999	.9992	.9963	.9876	.9662	.9231
	10	1.0000	1.0000	1.0000	1.0000	.9999	.9993	.9972	.9907	.9745
	11	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9995	.9981	.9937
	12	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9997	.9989
	13	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
16	0	.4401	.1853	.0743	.0281	.0100	.0033	.0010	.0003	.0001
	1	.8108	.5147	.2839	.1407	.0635	.0261	.0098	.0033	.0010
	2	.9571	.7892	.5614	.3518	.1971	.0994	.0451	.0183	.0066
	3	.9930	.9316	.7899	.5981	.4050	.2459	.1339	.0651	.0281
	4	.9991	.9830	.9209	.7982	.6302	.4499	.2892	.1666	.0853
	5	.9999	.9967	.9765	.9183	.8103	.6598	.4900	.3288	.1976
	6	1.0000	.9995	.9944	.9733	.9204	.8247	.6881	.5272	.3660
	7	1.0000	.9999	.9989	.9930	.9729	.9256	.8406	.7161	.5629
	8	1.0000	1.0000	.9998	.9985	.9925	.9743	.9329	.8577	.7441
	9	1.0000	1.0000	1.0000	.9998	.9984	.9929	.9771	.9417	.8759
	10	1.0000	1.0000	1.0000	1.0000	.9997	.9984	.9938	.9809	.9514
	11	1.0000	1.0000	1.0000	1.0000	1.0000	.9997	.9987	.9951	.9851
	12	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9998	.9991	.9965
	13	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9994
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999
	15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	16	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

（续附表 I）

n	x	.50	.55	.60	.65	θ .70	.75	.80	.85	.90	.95
15	0	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	1	.0005	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	2	.0037	.0011	.0003	.0001	.0000	.0000	.0000	.0000	.0000	.0000
	3	.0176	.0063	.0019	.0005	.0001	.0000	.0000	.0000	.0000	.0000
	4	.0592	.0255	.0093	.0028	.0007	.0001	.0000	.0000	.0000	.0000
	5	.1509	.0769	.0338	.0124	.0037	.0008	.0001	.0000	.0000	.0000
	6	.3036	.1818	.0950	.0422	.0152	.0042	.0008	.0001	.0000	.0000
	7	.5000	.3465	.2131	.1132	.0500	.0173	.0042	.0006	.0000	.0000
	8	.6964	.5478	.3902	.2452	.1311	.0566	.0181	.0036	.0003	.0000
	9	.8491	.7392	.5968	.4357	.2784	.1484	.0611	.0168	.0022	.0001
	10	.9408	.8796	.7827	.6481	.4845	.3135	.1642	.0617	.0127	.0006
	11	.9824	.9576	.9095	.8273	.7031	.5387	.3518	.1773	.0556	.0055
	12	.9963	.9893	.9729	.9383	.8732	.7639	.6020	.3958	.1841	.0362
	13	.9995	.9983	.9948	.9858	.9647	.9198	.8329	.6814	.4510	.1710
	14	1.0000	.9999	.9995	.9984	.9953	.9866	.9648	.9126	.7941	.5367
	15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
16	0	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	1	.0003	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	2	.0021	.0006	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	3	.0106	.0035	.0009	.0002	.0000	.0000	.0000	.0000	.0000	.0000
	4	.0384	.0149	.0049	.0013	.0003	.0000	.0000	.0000	.0000	.0000
	5	.1051	.0486	.0191	.0062	.0016	.0003	.0000	.0000	.0000	.0000
	6	.2272	.1241	.0583	.0229	.0071	.0016	.0002	.0000	.0000	.0000
	7	.4018	.2559	.1423	.0671	.0257	.0075	.0015	.0002	.0000	.0000
	8	.5982	.4371	.2839	.1594	.0744	.0271	.0070	.0011	.0001	.0000
	9	.7728	.6340	.4728	.3119	.1753	.0796	.0267	.0056	.0005	.0000
	10	.8949	.8024	.6712	.5100	.3402	.1897	.0817	.0235	.0033	.0001
	11	.9616	.9147	.8334	.7108	.5501	.3698	.2018	.0791	.0170	.0009
	12	.9894	.9719	.9349	.8661	.7541	.5950	.4019	.2101	.0684	.0070
	13	.9979	.9934	.9817	.9549	.9006	.8729	.6482	.4386	.2108	.0429
	14	.9997	.9990	.9967	.9902	.9739	.9365	.8593	.7161	.4853	.1892
	15	1.0000	.9999	.9997	.9990	.9967	.9900	.9719	.9257	.8147	.5599
	16	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

（续附表 I ）

n	x	.05	.10	.15	.20	θ .25	.30	.35	.40	.45
17	0	.4181	.1668	.0631	.0225	.0075	.0023	.0007	.0002	.0000
	1	.7922	.4818	.2525	.1182	.0501	.0193	.0067	.0021	.0006
	2	.9497	.7618	.5198	.3096	.1637	.0774	.0327	.0123	.0041
	3	.9912	.9174	.7556	.5489	.3530	.2019	.1028	.0464	.0184
	4	.9988	.9779	.9013	.7582	.5739	.3887	.2348	.1260	.0596
	5	.9999	.9953	.9681	.8943	.7653	.5968	.4197	.2639	.1471
	6	1.0000	.9992	.9917	.9623	.8929	.7752	.6188	.4478	.2902
	7	1.0000	.9999	.9983	.9891	.9598	.8954	.7872	.6405	.4743
	8	1.0000	1.0000	.9997	.9974	.9876	.9597	.9006	.8011	.6626
	9	1.0000	1.0000	1.0000	.9995	.9969	.9873	.9617	.9081	.8166
	10	1.0000	1.0000	1.0000	.9999	.9994	.9968	.9880	.9652	.9174
	11	1.0000	1.0000	1.0000	1.0000	.9999	.9993	.9970	.9894	.9699
	12	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9994	.9975	.9914
	13	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9995	.9981
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9997
	15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	16	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	17	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
18	0	.3972	.1501	.0536	.0180	.0056	.0016	.0004	.0001	.0000
	1	.7735	.4503	.2241	.0991	.0395	.0142	.0046	.0013	.0003
	2	.9419	.7338	.4797	.2713	.1353	.0600	.0236	.0082	.0025
	3	.9891	.9018	.7202	.5010	.3057	.1646	.0783	.0328	.0120
	4	.9985	.9718	.8794	.7164	.5187	.3327	.1886	.0942	.0411
	5	.9998	.9936	.9581	.8671	.7175	.5344	.3550	.2088	.1077
	6	1.0000	.9988	.9882	.9487	.8610	.7217	.5491	.3743	.2258
	7	1.0000	.9998	.9973	.9837	.9431	.8593	.7283	.5634	.3915
	8	1.0000	1.0000	.9995	.9957	.9807	.9404	.8609	.7368	.5778
	9	1.0000	1.0000	.9999	.9991	.9946	.9790	.9403	.8653	.7473
	10	1.0000	1.0000	1.0000	.9998	.9988	.9939	.9788	.9424	.8720
	11	1.0000	1.0000	1.0000	1.0000	.9998	.9986	.9938	.9797	.9463
	12	1.0000	1.0000	1.0000	1.0000	1.0000	.9997	.9986	.9942	.9817
	13	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9997	.9987	.9951
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9998	.9990
	15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999
	16	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	17	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	18	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

(续附表Ⅱ)

n	x	θ .05	.10	.15	.20	.25	.30	.35	.40	.45
19	0	.3774	.1351	.0456	.0144	.0042	.0011	.0003	.0001	.0000
	1	.7547	.4203	.1985	.0829	.0310	.0104	.0031	.0008	.0002
	2	.9335	.7054	.4413	.2369	.1113	.0462	.0170	.0055	.0015
	3	.9868	.8850	.6841	.4551	.2631	.1332	.0591	.0230	.0077
	4	.9980	.9648	.8556	.6733	.4654	.2822	.1500	.0696	.0280
	5	.9998	.9914	.9463	.8369	.6678	.4739	.2968	.1629	.0777
	6	1.0000	.9983	.9837	.9324	.8251	.6655	.4812	.3081	.1727
	7	1.0000	.9997	.9959	.9767	.9225	.8180	.6656	.4878	.3169
	8	1.0000	1.0000	.9992	.9933	.9713	.9161	.8145	.6675	.4940
	9	1.0000	1.0000	.9999	.9984	.9911	.9674	.9125	.8139	.6710
	10	1.0000	1.0000	1.0000	.9997	.9977	.9895	.9653	.9115	.8159
	11	1.0000	1.0000	1.0000	1.0000	.9995	.9972	.9886	.9648	.9129
	12	1.0000	1.0000	1.0000	1.0000	.9999	.9994	.9969	.9884	.9658
	13	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9993	.9969	.9891
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9994	.9972
	15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9995
	16	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999
	17	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	18	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	19	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
20	0	.3585	.1216	.0388	.0115	.0032	.0008	.0002	.0000	.0000
	1	.7358	.3917	.1756	.0692	.0243	.0076	.0021	.0005	.0001
	2	.9245	.6769	.4049	.2061	.0913	.0355	.0121	.0036	.0009
	3	.9841	.8670	.6477	.4114	.2252	.1071	.0444	.0160	.0049
	4	.9974	.9568	.8298	.6296	.4148	.2375	.1182	.0510	.0189
	5	.9997	.9887	.9327	.8042	.6172	.4164	.2454	.1256	.0553
	6	1.0000	.9976	.9781	.9133	.7858	.6080	.4166	.2500	.1299
	7	1.0000	.9996	.9941	.9679	.8982	.7723	.6010	.4159	.2520
	8	1.0000	.9999	.9987	.9900	.9591	.8867	.7624	.5956	.4143
	9	1.0000	1.0000	.9998	.9974	.9861	.9520	.8782	.7553	.5914
	10	1.0000	1.0000	1.0000	.9994	.9961	.9829	.9468	.8725	.7507
	11	1.0000	1.0000	1.0000	.9999	.9991	.9949	.9804	.9435	.8692
	12	1.0000	1.0000	1.0000	1.0000	.9998	.9987	.9940	.9790	.9420
	13	1.0000	1.0000	1.0000	1.0000	1.0000	.9997	.9985	.9935	.9786
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9997	.9984	.9936
	15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9997	.9985
	16	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9997
	17	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	18	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	19	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	20	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

（续附表 I）

n	x					θ					
		.50	.55	.60	.65	.70	.75	.80	.85	.90	.95
19	0	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	1	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	2	.0004	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	3	.0022	.0005	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	4	.0096	.0028	.0006	.0001	.0000	.0000	.0000	.0000	.0000	.0000
	5	.0318	.0109	.0031	.0007	.0001	.0000	.0000	.0000	.0000	.0000
	6	.0835	.0342	.0116	.0031	.0006	.0001	.0000	.0000	.0000	.0000
	7	.1796	.0871	.0352	.0114	.0028	.0005	.0000	.0000	.0000	.0000
	8	.3238	.1841	.0885	.0347	.0105	.0023	.0003	.0000	.0000	.0000
	9	.5000	.3290	.1861	.0875	.0326	.0089	.0016	.0001	.0000	.0000
	10	.6762	.5060	.3325	.1855	.0839	.0287	.0067	.0008	.0000	.0000
	11	.8204	.6831	.5122	.3344	.1820	.0775	.0233	.0041	.0003	.0000
	12	.9165	.8273	.6919	.5188	.3345	.1749	.0676	.0163	.0017	.0000
	13	.9682	.9223	.8371	.7032	.5261	.3322	.1631	.0537	.0086	.0002
	14	.9904	.9720	.9304	.8500	.7178	.5346	.3267	.1444	.0352	.0020
	15	.9978	.9923	.9770	.9409	.8668	.7369	.5449	.3159	.1150	.0132
	16	.9996	.9985	.9945	.9830	.9538	.8887	.7631	.5587	.2946	.0665
	17	1.0000	.9998	.9992	.9969	.9896	.9690	.9171	.8015	.5797	.2453
	18	1.0000	1.0000	.9999	.9997	.9989	.9958	.9856	.9544	.8649	.6226
	19	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
20	0	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	1	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	2	.0002	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	3	.0013	.0003	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	4	.0059	.0015	.0003	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	5	.0207	.0064	.0016	.0003	.0000	.0000	.0000	.0000	.0000	.0000
	6	.0577	.0214	.0065	.0015	.0003	.0000	.0000	.0000	.0000	.0000
	7	.1316	.0580	.0210	.0060	.0013	.0002	.0000	.0000	.0000	.0000
	8	.2517	.1308	.0565	.0196	.0051	.0009	.0001	.0000	.0000	.0000
	9	.4119	.2493	.1275	.0532	.0171	.0039	.0006	.0000	.0000	.0000
	10	.5881	.4086	.2447	.1218	.0480	.0139	.0026	.0002	.0000	.0000
	11	.7483	.5857	.4044	.2376	.1133	.0409	.0100	.0013	.0001	.0000
	12	.8684	.7480	.5841	.3990	.2277	.1018	.0321	.0059	.0004	.0000
	13	.9423	.8701	.7500	.5834	.3920	.2142	.0867	.0219	.0024	.0000
	14	.9793	.9447	.8744	.7546	.5836	.3828	.1958	.0673	.0113	.0003
	15	.9941	.9811	.9490	.8818	.7625	.5852	.3704	.1702	.0432	.0026
	16	.9987	.9951	.9840	.9556	.8929	.7748	.5886	.3523	.1330	.0159
	17	.9998	.9991	.9964	.9879	.9645	.9087	.7939	.5951	.3231	.0755
	18	1.0000	.9999	.9995	.9979	.9924	.9757	.9308	.8244	.6083	.2642
	19	1.0000	1.0000	1.0000	.9998	.9992	.9968	.9885	.9612	.8784	.6415
	20	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

附表Ⅲ：单样本 K-S 检验统计量

双侧检验的右尾概率

N	.200	.100	.050	.020	.010
1	.900	.950	.975	.990	.995
2	.684	.776	.842	.900	.929
3	.565	.636	.708	.785	.829
4	.493	.565	.624	.689	.734
5	.447	.509	.563	.627	.669
6	.410	.468	.519	.577	.617
7	.381	.436	.483	.538	.576
8	.358	.410	.454	.507	.542
9	.339	.387	.430	.480	.513
10	.323	.369	.409	.457	.489
11	.308	.352	.391	.437	.468
12	.296	.338	.375	.419	.449
13	.285	.325	.361	.404	.432
14	.275	.314	.349	.390	.418
15	.266	.304	.338	.377	.404
16	.258	.295	.327	.366	.392
17	.250	.286	.318	.355	.381
18	.244	.279	.309	.346	.371
19	.237	.271	.310	.337	.361
20	.232	.265	.294	.329	.352
	.100	.050	.025	.010	.005

N	.200	.100	.050	.020	.010
21	.226	.259	.287	.321	.344
22	.221	.253	.281	.314	.337
23	.216	.247	.275	.307	.330
24	.212	.242	.269	.301	.323
25	.208	.238	.264	.295	.317
26	.204	.233	.259	.290	.311
27	.200	.229	.254	.284	.305
28	.197	.225	.250	.279	.300
29	.193	.221	.246	.275	.295
30	.190	.218	.242	.270	.290
31	.187	.214	.238	.266	.285
32	.184	.211	.234	.262	.281
33	.182	.208	.231	.258	.277
34	.179	.205	.227	.254	.273
35	.177	.202	.224	.251	.269
36	.174	.199	.221	.247	.265
37	.172	.196	.218	.244	.262
38	.170	.194	.215	.241	.258
39	.168	.191	.213	.238	.255
40	.165	.189	.210	.235	.252
	.100	.050	.025	.010	.005

单侧检验的右尾概率

如果 $N>40$，则按下面的计算得到近似的概率：

双侧检验的右尾概率

.200	.100	.050	.020	.010
$1.07\sqrt{N}$	$1.22\sqrt{N}$	$1.36\sqrt{N}$	$1.52\sqrt{N}$	$1.63\sqrt{N}$
.100	.050	.025	.010	.005

单侧检验的右尾概率

附表Ⅳ：正态分布表

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641
0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
0.9	.1814	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
3.5	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002

(续附表V)

Z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9614	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.8950	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990

Z	1.282	1.645	1.960	2.326	2.576
$1-\phi(Z)$	0.100	0.050	0.025	0.010	0.005

附表Ⅵ：带有 $Q=0.05$ 的累积二项分布表（用于符号检验）

n	左 S	P	右 S	n	左 S	P	右 S	n	左 S	P	右 S
1	0	.5000	1		1	.0195	8		2	.0112	11
2	0	.2500	2		2	.0898	7		3	.0461	10
	1	.7500	1		3	.2539	6		4	.1334	9
3	0	.1250	3		4	.5000	5		5	.2905	8
	1	.5000	2	10	0	.0010	10		6	.5000	7
4	0	.0625	4		1	.0107	9	14	0	.0000	14
	1	.3125	3		2	.0547	8		1	.0009	13
	2	.6875	2		3	.1719	7		2	.0065	12
5	0	.0312	5		4	.3770	6		3	.0287	11
	1	.1875	4		5	.6230	5		4	.0898	10
	2	.5000	3	11	0	.0005	11		5	.2120	9
6	0	.0156	6		1	.0059	10		6	.3953	8
	1	.1094	5		2	.0327	9		7	.6047	7
	2	.3438	4		3	.1133	8	15	0	.0000	15
	3	.6562	3		4	.2744	7		1	.0005	14
7	0	.0078	7		5	.5000	6		2	.0037	13
	1	.0625	6	12	0	.0002	12		3	.0176	12
	2	.2266	5		1	.0032	11		4	.0592	11
	3	.5000	4		2	.0193	10		5	.1509	10
8	0	.0039	8		3	.0730	9		6	.3036	9
	1	.0352	7		4	.1938	8		7	.5000	8
	2	.1445	6		5	.3872	7	16	0	.0000	16
	3	.3633	5		6	.6128	6		1	.0003	15
	4	.6367	4	13	0	.0001	13		2	.0021	14
9	0	.0020	9		1	.0017	12		3	.0106	13

(续附表Ⅵ)

n	左 S	P	右 S	n	左 S	P	右 S	n	左 S	P	右 S
	4	.0384	12		1	.0001	17		6	.0835	13
	5	.1051	11		2	.0007	16		7	.1796	12
	6	.2272	10		3	.0038	15		8	.3238	11
	7	.4018	9		4	.0154	14		9	.5000	10
	8	.5982	8		5	.0481	13	20	0	.0000	20
17	0	.0000	17		6	.1189	12		1	.0000	19
	1	.0001	16		7	.2403	11		2	.0002	18
	2	.0012	15		8	.4073	10		3	.0013	17
	3	.0064	14		9	.5927	9		4	.0059	16
	4	.0245	13	19	0	.0000	19		5	.0207	15
	5	.0717	12		1	.0000	18		6	.0577	14
	6	.1662	11		2	.0004	17		7	.1316	13
	7	.3145	10		3	.0022	16		8	.2517	12
	8	.5000	9		4	.0096	15		9	.4119	11
18	0	.0000	18		5	.0318	14		10	.5881	10

如果 $n>20$，则按下面公式计算在附表Ⅳ中得到概率：

$$z_{+,R}=\frac{S_{+}-0.5-.5n}{.5\sqrt{n}}, \qquad z_{-,R}=\frac{S_{-}-0.5-.5n}{.5\sqrt{n}},$$

预定的	近似的
S_+的右尾概率	$Z_{+,R}$的右尾概率
S_-的右尾概率	$Z_{-,R}$的右尾概率

附表Ⅶ：Wilcoxon 符号秩检验统计量

n	左 T	P	右 T	n	左 T	P	右 T	n	左 T	P	右 T
2	0	.250	3		8	.344	13		10	.156	26
	1	.500	2		9	.422	12		11	.191	25
3	0	.125	6		10	.500	11		12	.230	24
	1	.250	5	7	0	.008	28		13	.273	23
	2	.375	4		1	.016	27		14	.320	22
	3	.625	3		2	.023	26		15	.371	21
4	0	.062	10		3	.039	25		16	.422	20
	1	.125	9		4	.055	24		17	.473	19
	2	.188	8		5	.078	23		18	.527	18
	3	.312	7		6	.109	22	9	0	.002	45
	4	.438	6		7	.148	21		1	.004	44
	5	.562	5		8	.188	20		2	.006	43
5	0	.031	15		9	.234	19		3	.010	42
	1	.062	14		10	.289	18		4	.014	41
	2	.094	13		11	.344	17		5	.020	40
	3	.156	12		12	.406	16		6	.027	39
	4	.219	11		13	.469	15		7	.037	38
	5	.312	10		14	.531	14		8	.049	37
	6	.406	9	8	0	.004	36		9	.064	36
	7	.500	8		1	.008	35		10	.082	35
6	0	.016	21		2	.012	34		11	.102	34
	1	.031	20		3	.020	33		12	.125	33
	2	.047	19		4	.027	32		13	.150	32
	3	.078	18		5	.039	31		14	.180	31
	4	.109	17		6	.055	30		15	.213	30
	5	.156	16		7	.074	29		16	.248	29
	6	.219	15		8	.098	28		17	.285	28
	7	.281	14		9	.125	27		18	.326	27

（续附表Ⅶ）

n	左 T	P	右 T	n	左 T	P	右 T	n	左 T	P	右 T
	19	.367	36		26	.461	29		28	.350	38
	20	.410	25		27	.500	28		29	.382	37
	21	.455	24	11	0	.000	66		30	.416	36
	22	.500	23		1	.001	65		31	.449	35
10	0	.001	55		2	.001	64		32	.483	34
	1	.002	54		3	.002	63		33	.517	33
	2	.003	53		4	.003	62	12	0	.000	78
	3	.005	52		5	.005	61		1	.000	77
	4	.007	51		6	.007	60		2	.001	76
	5	.010	50		7	.009	59		3	.001	75
	6	.014	49		8	.012	58		4	.002	74
	7	.019	48		9	.016	57		5	.002	73
	8	.024	47		10	.021	56		6	.003	72
	9	.032	46		11	.027	55		7	.005	71
	10	.042	45		12	.034	54		8	.006	70
10	11	.053	44		13	.042	53		9	.008	69
	12	.065	43		14	.051	52		10	.010	68
	13	.080	42		15	.062	51		11	.013	67
	14	.097	41		16	.074	50		12	.017	66
	15	.116	40		17	.087	49		13	.021	65
	16	.138	39		18	.103	48		14	.026	64
	17	.161	38		19	.120	47		15	.032	63
	18	.188	37		20	.139	46		16	.039	62
	19	.216	36		21	.160	45		17	.046	61
	20	.246	35		22	.183	44		18	.055	60
	21	.278	34		23	.207	43		19	.065	59
	22	.312	33		24	.232	42		20	.076	58
	23	.348	32		25	.260	41		21	.088	57
	24	.385	31		26	.289	40		22	.102	56
	25	.423	30		27	.319	39		23	.117	55

(续附表Ⅶ)

n	左 T	P	右 T	n	左 T	P	右 T	n	左 T	P	右 T
14	16	.010	89		49	.428	56		29	.042	91
	17	.012	88		50	.452	55		30	.047	90
	18	.015	87		51	.476	54	15	31	.053	89
	19	.018	86		52	.500	53		32	.060	88
	20	.021	85	15	0	.000	120		33	.068	87
	21	.025	84		1	.000	119		34	.076	86
	22	.029	83		2	.000	118		35	.084	85
	23	.034	82		3	.000	117		36	.094	84
	24	.039	81		4	.000	116		37	.104	83
	25	.045	80		5	.000	115		38	.115	82
	26	.052	79		6	.000	114		39	.126	81
	27	.059	78		7	.001	113		40	.138	80
	28	.068	77		8	.001	112		41	.151	79
	29	.077	76		9	.001	111		42	.165	78
	30	.086	75		10	.001	110		43	.180	77
	31	.097	74		11	.002	109		44	.195	76
	32	.108	73		12	.002	108		45	.211	75
	33	.121	72		13	.003	107		46	.227	74
	34	.134	71		14	.003	106		47	.244	73
	35	.148	70		15	.004	105		48	.262	72
	36	.163	69		16	.005	104		49	.281	71
	37	.179	68		17	.006	103		50	.300	70
	38	.196	67		18	.008	102		51	.319	69
	39	.213	66		19	.009	101		52	.339	68
	40	.232	65		20	.011	100		53	.360	67
	41	.251	64		21	.013	99		54	.381	66
	42	.271	63		22	.015	98		55	.402	65
	43	.292	62		23	.018	97		56	.423	64
	44	.313	61		24	.021	96		57	.445	63
	45	.335	60		25	.024	95		58	.467	62
	46	.357	59		26	.028	94		59	.489	61
	47	.380	58		27	.032	93		60	.511	60
	48	.404	57		28	.036	92				

如果 $n>15$，则按下式，在附表Ⅳ中查找概率

$$z_{+,R}=\frac{T_{+}-0.5-n(n+1)/4}{\sqrt{n(n+1)(2n+1)/24}} \qquad z_{-,R}=\frac{T_{-}-0.5-n(n+1)/4}{\sqrt{n(n+1)(2n+1)/24}}$$

预定的	近似的
T_+的右尾概率	$z_{+,R}$的右尾概率
T_-的右尾概率	$z_{-,R}$的右尾概率

左尾概率

m	n	U	P	m	n	U	P	m	n	U	P	m	n	U	P
2	2	2	.333	2	18	2	.011			3	.025	4	10	2	.002
2	3	2	.200			3	.105			4	.101			3	.014
		3	.500			4	.284			5	.350			4	.068
2	4	2	.133	3	3	2	.100	3	15	2	.002			5	.203
		3	.400			3	.300			3	.022			6	.419
2	5	2	.095	3	4	2	.057			4	.091	4	11	2	.001
		3	.333			3	.200			5	.331			3	.011
2	6	2	.071	3	5	2	.036	3	16	2	.022			4	.055
		3	.286			3	.143			3	.020			5	.176
2	7	2	.056			4	.429			4	.082			6	.374
		3	.250	3	6	2	.024			5	.314	4	12	2	.001
2	8	2	.044			3	.107	3	17	2	.002			3	.009
		3	.222			4	.345			3	.018			4	.045
2	9	2	.036	3	7	2	.017			4	.074			5	.154
		3	.200			3	.083			5	.298			6	.335
		4	.491			4	.283	4	4	2	.029	4	13	2	.001
2	10	2	.030	3	8	2	.012			3	.114			3	.007
		3	.182			3	.067			4	.371			4	.037
		4	.455			4	.236	4	5	2	.016			5	.136
2	11	2	.026	3	9	2	.009			3	.071			6	.302
		3	.167			3	.055			4	.262	4	14	2	.001
		4	.423			4	.200			5	.500			3	.006
2	12	2	.022			5	.491	4	6	2	.010			4	.031
		3	.154	3	10	2	.007			3	.048			5	.121
		4	.396			3	.045			4	.190			6	.274
2	13	2	.019			4	.171			5	.405	4	15	2	.001
		3	.143			5	.455	4	7	2	.006			3	.005
		4	.371	3	11	2	.005			3	.033			4	.027
2	14	2	.017			3	.038			4	.142			5	.108
		3	.133			4	.148			5	.333			6	.249
		4	.350			5	.423	4	8	2	.004	4	16	2	.000
2	15	2	.015	3	12	2	.004			3	.024			3	.004
		3	.125			3	.033			4	.109			4	.023
		4	.331			4	.130			5	.279			5	.097
2	16	2	.013			5	.396	4	9	2	.003			6	.227
		3	.118	3	13	2	.004			3	.018	5	5	2	.008
		4	.314			3	.029			4	.085			3	.040
2	17	2	.012			4	.114			5	.236			4	.167
		3	.111			5	.371			6	.471			5	.357
		4	.298	3	14	2	.003								

左尾概率

m	n	U	P
5	6	2	.004
		3	.024
		4	.110
		5	.262
5	7	2	.003
		3	.015
		4	.076
		5	.197
		6	.424
5	8	2	.002
		3	.010
		4	.054
		5	.152
		6	.347
5	9	2	.001
		3	.007
		4	.039
		5	.119
		6	.287
5	10	2	.001
		3	.005
		4	.029
		5	.095
		6	.239
		7	.455
5	11	2	.000
		3	.004
		4	.022
		5	.077
		6	.201
		7	.407
5	12	2	.000
		3	.003
		4	.017
		5	.063
		6	.170
		7	.365
5	13	2	.000
		3	.002
		4	.013
		5	.053
		6	.145
		7	.330

m	n	U	P
5	14	2	.000
		3	.002
		4	.011
		5	.044
		6	.125
		7	.299
		8	.496
5	15	2	.000
		3	.001
		4	.009
		5	.037
		6	.108
		7	.272
		8	.460
6	6	2	.002
		3	.013
		4	.067
		5	.175
		6	.392
6	7	2	.001
		3	.008
		4	.043
		5	.121
		6	.296
		7	.500
6	8	2	.001
		3	.005
		4	.028
		5	.086
		6	.226
		7	.413
6	9	2	.000
		3	.003
		4	.019
		5	.063
		6	.175
		7	.343
6	10	2	.000
		3	.002
		4	.013
		5	.047
		6	.137
		7	.287
		8	.497

m	n	U	P
6	11	2	.000
		3	.001
		4	.009
		5	.036
		6	.108
		7	.242
		8	.436
6	12	2	.000
		3	.001
		4	.007
		5	.028
		6	.087
		7	.205
		8	.383
6	13	2	000
		3	.001
		4	.005
		5	.022
		6	.070
		7	.176
		8	.338
6	14	2	.000
		3	.001
		4	.004
		5	.017
		6	.058
		7	.151
		8	.299
7	7	2	.001
		3	.004
		4	.025
		5	.078
		6	.209
		7	.383
7	8	2	.000
		3	.002
		4	.015
		5	.051
		6	.149
		7	.296

m	n	U	P
7	9	2	.000
		3	.001
		4	.010
		5	.035
		6	.108
		7	.231
		8	.427
7	10	2	.000
		3	.001
		4	.006
		5	.024
		6	.080
		7	.182
		8	.355
7	11	2	.000
		3	.001
		4	.004
		5	.018
		6	.060
		7	.145
		8	.296
		9	.484
7	12	2	.000
		3	.000
		4	.003
		5	.013
		6	.046
		7	.117
		8	.247
		9	.428
7	13	2	.000
		3	.000
		4	.002
		5	.010
		6	.035
		7	.095
		8	.208
		9	.378
8	8	2	.000
		3	.001
		4	.009
		5	.032
		6	.100
		7	.214
		8	.405

（续附表Ⅷ）

左尾概率

m	n	U	P	m	n	U	P	m	n	U	P	m	n	U	P
8	9	2	.000			5	.012			7	.051			4	.000
		3	.001			6	.044			8	.128			5	.001
		4	.005			7	.109			9	.242			6	.005
		5	.020			8	.238			10	.414			7	.015
		6	.069			9	.399	10	11	2	.000			8	.044
		7	.157	9	10	2	.000			3	.000			9	.099
		8	.319			3	.000			4	.001			10	.202
		9	.500			4	.002			5	.003			11	.335
8	10	2	.000			5	.008			6	.012	12	12	2	.000
		3	.000			6	.029			7	.035			3	.000
		4	.003			7	.077			8	.092			4	.000
		5	.013			8	.179			9	.185			5	.001
		6	.048			9	.319			10	.335			6	.003
		7	.117	9	11	2	.000			11	.500			7	.009
		8	.251			3	.000	10	12	2	.000			8	.030
		9	.419			4	.001			3	.000			9	.070
8	11	2	.000			5	.005			4	.000			10	.150
		3	.000			6	.020			5	.002			11	.263
		4	.002			7	.055			6	.008			12	421
		5	.009			8	.135			7	.024	2	2	4	.333
		6	.034			9	.255			8	.067	2	3	5	.100
		7	.088			10	.430			9	.142			4	.500
		8	.199	9	12	2	.000			10	.271	2	4	5	.200
		9	.352			3	.000			11	.425	2	5	5	.286
8	12	2	.000			4	.001	11	11	2	000	2	6	5	.357
		3	.000			5	.003			3	.000	2	7	5	.417
		4	.001			6	.014			4	.000	2	8	5	.467
		5	.006			7	.040			5	.002	3	3	6	.100
		6	.025			8	.103			6	007			5	.300
		7	.067			9	.205			7	.023	3	4	7	.029
		8	.159			10	.362			8	.063			6	.200
		9	.297	10	10	2	.000			9	.135			5	.457
		10	.480			3	.000			10	.260	3	5	7	.071
9	9	2	.000			4	.001			11	.410			6	.286
		3	.000			5	.004	11	12	2	.000	3	6	7	.119
		4	.003			6	.019			3	.000			6	.357

(续附表Ⅶ)

右尾概率															
m	*n*	*U*	*P*	*m*	*n*	*U*	*P*	*m*	*n*	*U*	*P*	*m*	*n*	*U*	*P*
3	7	7	.167			8	.393			10	208			10	.395
		6	.417	4	14	9	.234			9	.465	6	13	13	.034
3	8	7	.212			8	.421	5	14	11	.111			12	.092
		6	.467	4	15	9	.258			10	.234			11	.257
3	9	7	.255			8	.446	5	15	11	.129			10	.439
3	10	7	.294	4	16	9	.282			10	.258	6	14	13	.044
3	11	7	.330			8	.470	6	6	12	.002			12	.111
3	12	7	.363	5	5	10	.008			11	.013			11	.295
3	13	7	.393			9	.040			10	.067			10	.480
3	14	7	.421			8	.167			9	.175	7	7	14	.001
3	15	7	.446			7	.357			8	.392			13	.004
3	16	7	.470	5	6	11	.002	6	7	13	.001			12	.025
3	17	7	.491			10	.024			12	.008			11	.078
4	4	8	.029			9	.089			11	.034			10	.209
		7	.114			8	.262			10	.121			9	.383
		6	.371			7	.478			9	.267	7	8	15	.000
4	5	9	.008	5	7	11	.008			8	.500			14	.002
		8	.071			10	.045	6	8	13	.002			13	.012
		7	.214			9	.146			12	.016			12	.051
		6	.500			8	.348			11	.063			11	.133
4	6	9	.024	5	8	11	.016			10	.179			10	.296
		8	.119			10	.071			9	.354			9	.486
		7	.310			9	.207	6	9	13	.006	7	9	15	.001
4	7	9	.045			8	.424			12	.028			14	.006
		8	.167	5	9	11	.028			11	.098			13	.025
		7	.394			10	.098			10	.238			12	.084
4	8	9	.071			9	.266			9	.434			11	.194
		8	.212			8	.490	6	10	13	.010			10	.378
		7	.467	5	10	11	.042			12	.042	7	10	15	.002
4	9	9	.098			10	.126			11	.136			14	.010
		8	.255			9	.322			10	.294			13	.043
4	10	9	.126	5	11	11	.058	6	11	13	.017			12	.121
		8	.294			10	.154			12	.058			11	.257
4	11	9	.154			9	.374			11	.176			10	.451
		8	.330	5	12	11	.075			10	346	7	11	15	.004
4	12	9	.181			10	.181	6	12	13	025			14	.017
		8	.363			9	.421			12	075			13	.064
4	13	9	.208	5	13	11	.092			11	217			12	.160

(续附表Ⅷ)

m	n	U	P	m	n	U	P	m	n	U	P	m	n	U	P
		11	.318	8	11	17	.001			16	.015			13	.395
7	12	15	.007			16	.004			15	.045	11	11	22	.000
		14	.025			15	.018			14	.115			21	.000
		13	.089			14	.057			13	.227			20	.000
		12	.199			13	.138			12	.395			19	.002
		11	.376			12	.278	10	10	20	.000			18	.007
7	13	15	.010			11	.453			19	.000			17	.023
		14	.034	8	12	17	.001			18	.000			16	.063
		13	.116			16	.007			17	.001			15	.135
		12	.238			15	.029			16	.004			14	.260
		11	.430			14	.080			17	.019			13	.410
8	8	16	000			13	.183			16	.051	11	12	23	.000
		15	.001			12	.337			15	.128			22	.000
		14	.009	9	9	18	.000			14	.242			21	.000
		13	.032			17	.000			13	.414			20	.001
		12	.100			16	.003	10	11	21	.000			19	.004
		11	.214			15	.012			20	.000			18	.015
		10	.405			14	044			19	.000			17	.041
8	9	17	.000			13	.109			18	.003			16	099
		16	.001			12	.238			17	.010			15	.191
		15	.004			11	.399			16	.035			14	.335
		14	.020	9	10	19	.000			15	.085			13	.493
		13	.061			18	.000			14	.185	12	12	24	.000
		12	.157			17	.001			13	.320			23	.000
		11	.298			16	.008			12	.500			22	.000
		10	.500			15	.026	10	12	21	.000			21	.001
8	10	17	.000			14	.077			20	.000			20	.003
		16	.002			13	.166			19	.001			19	.009
		15	.010			12	.319			18	.006			18	.030
		14	.036			11	.490			17	.020			17	.070
		13	.097	9	11	19	.000			16	.056			16	.150
		12	.218			18	.001			15	.125			15	.263
		11	.379			17	.003			14	.245			14	.421

如果 $m+n=N>20$ 和 $m>12$，$n>12$，则按下面的公式计算，在附表Ⅳ中得到相应的概率 P。

$$z_L=\frac{U+0.5-1-2mn/N}{\sqrt{\frac{2mn\ (2mn-N)}{N^2\ (N-1)}}}\qquad z_R=\frac{U-0.5-1-2mn/N}{\sqrt{\frac{2mn\ (2mn-N)}{N^2\ (N-1)}}}$$

预定的	近似的
U 的左尾概率	z_L 的左尾概率
U 的右尾概率	Z_R 的右尾概率

（续附表 IX）

N	V	左尾 P	V	右尾 P	N	V	左尾 P	V	右尾 P
24	1	.0000			25	1	.0000		
	2	.0000				2	.0000		
	3	.0000				3	.0000		
	4	.0000				4	.0000		
	5	.0000				5	.0000		
	6	.0000				6	.0000		
	7	.0000				7	.0000	24	.0000
	8	.0001	23	.0000		8	.0000	23	.0005
	9	.0008	22	.0007		9	.0003	22	.0037
	10	.0044	21	.0053		10	.0018	21	.0170
	11	.0177	20	.0235		11	.0084	20	.0564
	12	.0554	19	.0742		12	.0294	19	.1423
	13	.1374	18	.1783		13	.0815	18	.2852
	14	.2768	17	.3405		14	.1827	17	.4708
	15	.4631	16	.5369		15	.3384	16	.6616

如果 $N>25$，则按下面公式计算，在附表 IV 中查找相应的概率：

$$Z_L=\frac{V+0.5-(2N-1)/3}{\sqrt{(16N-29)/90}} \qquad Z_R=\frac{V-0.5-(2N-1)/3}{\sqrt{(16N-29)/90}}$$

预定的	近似的
V 的左尾概率	Z_L 的左尾概率
V 的右尾概率	Z_R 的右尾概率

附表 X：*Mann-Whitney-Wilcoxon* 分布表

$m=1$

n	左 T_x	P	右 T_x
1	1	.500	2
2	1	.333	3
	2	.667	2
3	1	.250	4
	2	.500	3
4	1	.200	5
	2	.400	4
	3	.600	3
5	1	.167	6
	2	.333	5
	3	.500	4
6	1	.143	7
	2	.286	6
	3	.429	5
	4	.571	4
7	1	.125	8
	2	.250	7
	3	.375	6
	4	.500	5
8	1	.111	9
	2	.222	8
	3	.333	7
	4	.444	6
	5	.556	5
9	1	.100	10
	2	.200	9
	3	.300	8
	4	.400	7
	5	.500	6
10	1	.091	11
	2	.182	10
	3	.273	9
	4	.364	8
	5	.455	7
	6	.545	6

$m=2$

n	左 T_x	P	
2	3	.167	7
	4	.333	6
	5	.667	5
3	3	.100	9
	4	.200	8
	5	.400	7
	6	.600	6
4	3	.067	11
	4	.133	10
	5	.267	9
	6	.400	8
	7	.600	7
5	3	.048	13
	4	.095	12
	5	.190	11
	6	.286	10
	7	.429	9
	8	.571	8
6	3	.036	15
	4	.071	14
	5	.143	13
	6	.214	12
	7	.321	11
	8	.429	10
	9	.571	9
7	3	.028	17
	4	.056	16
	5	.111	15
	6	.167	14
	7	.250	13
	8	.333	12
	9	.444	11
	10	.556	10

$m=2$

n	左 T_x	P	右 T_x
8	3	.022	19
	4	.044	18
	5	.089	17
	6	.133	16
	7	.200	15
	8	.267	14
	9	.356	13
	10	.444	12
	11	.556	11
9	3	.018	21
	4	.036	20
	5	.073	19
	6	.109	18
	7	.164	17
	8	.218	16
	9	.291	15
	10	.364	14
	11	.455	13
	12	.545	12
10	3	.015	23
	4	.030	22
	5	.061	21
	6	.091	20
	7	.136	19
	8	.182	18
	9	.242	17
	10	.303	16
	11	.379	15
	12	.455	14
	13	.545	13

（续附表 X）

n	左 T_x	P	右 T_x
	$m=4$		
7	10	.003	38
	11	.006	37
	12	.012	36
	13	.021	35
	14	.036	34
	15	.055	33
	16	.082	32
	17	.115	31
	18	.158	30
	19	.206	29
	20	.264	28
	21	.324	27
	22	.394	26
	23	.464	25
	24	.536	24
8	10	.002	42
	11	.004	41
	12	.008	40
	13	.014	39
	14	.024	38
	15	.036	37
	16	.055	36
	17	.077	35
	18	.107	34
	19	.141	33
	20	.184	32
	21	.230	31
	22	.285	30
	23	.341	29
	24	.404	28
	25	.467	27
	26	.533	26

n	左 T_x	P	右 T_x
	$m=4$		
9	10	.001	46
	11	.003	45
	12	.006	44
	13	.010	43
	14	.017	42
	15	.025	41
	16	.038	40
	17	.053	39
	18	.074	38
	19	.099	37
	20	.130	36
	21	.165	35
	22	.207	34
	23	.252	33
	24	.302	32
	25	.355	31
	26	.413	30
	27	.470	29
	28	.530	28
10	10	.001	50
	11	.002	49
	12	.004	48
	13	.007	47
	14	.012	46
	15	.018	45
	16	.027	44
	17	.038	43
	18	.053	42
	19	.071	41
	20	.094	40
	21	.120	39
	22	.152	38
	23	.187	37
	24	.227	36
	25	.270	35
	26	.318	34
	27	.367	33
	28	.420	32
	29	.473	31
	30	.527	30

n	左 T_x	P	右 T_x
	$m=5$		
5	15	.004	40
	16	.008	39
	17	.016	38
	18	.028	37
	19	.048	36
	20	.075	35
	21	.111	34
	22	.155	33
	23	.210	32
	24	.274	31
	25	.345	30
	26	.421	29
	27	.500	28
6	15	.002	45
	16	.004	44
	17	.009	43
	18	.015	42
	19	.026	41
	20	.041	40
	21	.063	39
	22	.089	38
	23	.123	37
	24	.165	36
	25	.214	35
	26	.268	34
	27	.331	33
	28	.396	32
	29	.465	31
	30	.535	30

(续附表 X)

n	左 T_x	P	右 T_x	n	左 T_x	P	右 T_x	n	左 T_x	P	右 T_x
		$m=5$				$m=5$				$m=5$	
7	15	.001	50		23	.047	47		28	.120	47
	16	.003	49		24	.064	46		29	.149	46
	17	.005	48		25	.085	45		30	.182	45
	18	.009	47		26	.111	44		31	.219	44
	19	.015	46		27	.142	43		32	.259	43
	20	.024	45		28	.177	42		33	.303	42
	21	.037	44		29	.218	41		34	.350	41
	22	.053	43		30	.262	40		35	.399	40
	23	.074	42		31	.311	39		36	.449	39
	24	.101	41		32	.362	38		37	.500	38
	25	.134	40		33	.416	37	10	15	.000	65
	26	.172	39		34	.472	36		16	.001	64
	27	.216	38		35	.528	35		17	.001	63
	28	.265	37	9	15	.000	60		18	.002	62
	29	.319	36		16	.001	59		19	.004	61
	30	.378	35		17	.002	58		20	.006	60
	31	.438	34		18	.003	57		21	.010	59
	32	.500	33		19	.006	56		22	.014	58
8	15	.001	55		20	.009	55		23	.020	57
	16	.002	54		21	.014	54		24	.028	56
	17	.003	53		22	.021	53		25	.038	55
	18	.005	52		23	.030	52		26	.050	54
	19	.009	51		24	.041	51		27	.065	53
	20	.015	50		25	.056	50		28	.082	52
	21	.023	49		26	.073	49		29	.103	51
	22	.033	48		27	.095	48		30	.127	50

（续附表 X）

n	左 T_x	P	右 T_x	n	左 T_x	P	右 T_x	n	左 T_x	P	右 T_x
	$m=6$				$m=6$				$m=6$		
8	32	.054	58		33	.044	63		31	.016	71
	33	.071	57		34	.057	62		32	.021	70
	34	.091	56		35	.072	61		33	.028	69
	35	.114	55		36	.091	60		34	.036	68
	36	.141	54		37	.112	59		35	.047	67
	37	.172	53		38	.136	58		36	.059	66
	38	.207	52		39	.164	57		37	.074	65
	39	.245	51		40	.194	56		38	.090	64
	40	.286	50	9	41	.228	55		39	.110	63
	41	.331	49		42	.264	54		40	.132	62
	42	.377	48		43	.303	53		41	.157	61
	43	.426	47		44	.344	52		42	.184	60
	44	.475	46		45	.388	51		43	.214	59
	45	.525	45		46	.432	50		44	.246	58
9	21	.000	75		47	.477	49		45	.281	57
	22	.000	74		48	.523	48		46	.318	56
	23	.001	73	10	21	.000	81		47	.356	55
	24	.001	72		22	.000	80		48	.396	54
	25	.002	71		23	.000	79		49	.437	53
	26	.004	70		24	.001	78		50	.479	52
	27	.006	69		25	.001	77		51	.521	51
	28	.009	68		26	.002	76				
	29	.013	67		27	.004	75				
	30	.018	66		28	.005	74				
	31	.025	65		29	.008	73				
	32	.033	64		30	.011	72				

(续附表 X)

n	左 T_x	P	右 T_x	n	左 T_x	P	右 T_x	n	左 T_x	P	右 T_x
		$m=7$				$m=7$				$m=7$	
7	28	.000	77		29	.000	83		55	.478	57
	29	.001	76		30	.001	82		56	.522	56
	30	.001	75		31	.001	81	9	28	.000	91
	31	.002	74		32	.002	80		29	.000	90
	32	.003	73		33	.003	79		30	.000	89
	33	.006	72		34	.005	78		31	.001	88
	34	.009	71		35	.007	77		32	.001	87
	35	.013	70		36	.010	76		33	.002	86
	36	.019	69		37	.014	75		34	.003	85
	37	.027	68		38	.020	74		35	.004	84
	38	.036	67		39	.027	73		36	.006	83
	39	.049	66		40	.036	72		37	.008	82
	40	.064	65		41	.047	71		38	.011	81
	41	.082	64		42	.060	70		39	.016	80
	42	.104	63		43	.076	69		40	.021	79
	43	.130	62		44	.095	68		41	.027	78
	44	.159	61		45	.116	67		42	.036	77
	45	.191	60		46	.140	66		43	.045	76
	46	.228	59		47	.168	65		44	.057	75
	47	.267	58	8	48	.198	64		45	.071	74
	48	.310	57		49	.232	63		46	.087	73
	49	.355	56		50	.268	62		47	.105	72
	50	.402	55		51	.306	61		48	.126	71
	51	.451	54		52	.347	60		49	.150	70
	52	.500	53		53	.389	59		50	.176	69
8	28	.000	84		54	.433	58		51	.204	68

(续附表X)

n	左 T_x	P	右 T_x	n	左 T_x	P	右 T_x	n	左 T_x	P	右 T_x
	$m=7$				$m=7$				$m=8$		
9	52	.235	67		46	.054	80	8	36	.000	100
	53	.268	66		47	.067	79		37	.000	99
	54	.303	65		48	.081	78		38	.000	98
	55	.340	64		49	.097	77		39	.001	97
	56	.379	63		50	.115	76		40	.001	96
	57	.419	62		51	.135	75		41	.001	95
	58	.459	61		52	.157	74		42	.002	94
	59	.500	60		53	.182	73		43	.003	93
10	28	.000	98		54	.209	72		44	.005	92
	29	.000	97		55	.237	71		45	.007	91
	30	.000	96		56	.268	70		46	.010	90
	31	.000	95		57	.300	69		47	.014	89
	32	.001	94		58	.335	68		48	.019	88
	33	.001	93		59	.370	67		49	.025	87
	34	.002	92		60	.406	66		50	.032	86
	35	.002	91		61	.443	65		51	.041	85
	36	.003	90		62	.481	64		52	.052	84
	37	.005	89		63	.519	63		53	.065	83
	38	.007	88						54	.080	82
	39	.009	87						55	.097	81
	40	.012	86						56	.117	80
	41	.017	85						57	.139	79
	42	.022	84						58	.164	78
	43	.028	83						59	.191	77
	44	.035	82						60	.221	76
	45	.044	81						61	.253	75

（续附表 X）

n	左 T_x	P	右 T_x	n	左 T_x	P	右 T_x	n	左 T_x	P	右 T_x
		$m=8$				$m=8$				$m=8$	
8	62	.287	74		55	.057	89		44	.002	108
	63	.323	73		56	.069	88		45	.002	107
	64	.360	72		57	.084	87		46	.003	106
	65	.399	71		58	.100	86		47	.004	105
	66	.439	70		59	.118	85		48	.006	104
	67	.480	69		60	.138	84		49	.008	103
	68	.520	68		61	.161	83		50	.010	102
9	36	.000	108		62	.185	82		51	.013	101
	37	.000	107		63	.212	81		52	.017	100
	38	.000	106		64	.240	80		53	.022	99
	39	.000	105		65	.271	79		54	.027	98
	40	.000	104		66	.303	78		55	.034	97
	41	.001	103		67	.336	77		56	.042	96
	42	.001	102		68	.371	76		57	.051	95
	43	.002	101		69	.407	75		58	.061	94
9	44	.003	100		70	.444	74		59	.073	93
	45	.004	99		71	.481	73		60	.086	92
	46	.006	98		72	.519	72		61	.102	91
	47	.008	97	10	36	.000	116		62	.118	90
	48	.010	96		37	.000	115		63	.137	89
	49	.014	95		38	.000	114		64	.158	88
	50	.018	94		39	.000	113		65	.180	87
	51	.023	93		40	.000	112		66	.204	86
	52	.030	92		41	.000	111		67	.230	85
	53	.037	91		42	.001	110		68	.257	84
	54	.046	90		43	.001	109		69	.286	83

（续附表X）

n	左 T_x	P	右 T_x	n	左 T_x	P	右 T_x	n	左 T_x	P	右 T_x
		$m=8$				$m=9$				$m=9$	
10	70	.317	82	9	45	.000	126		71	.111	100
	71	.348	81		46	.000	125		72	.129	99
	72	.381	80		47	.000	124		73	.149	98
	73	.414	79		48	.000	123		74	.170	97
	74	.448	78		49	.000	122		75	.193	96
	75	.483	77		50	.000	121		76	.218	95
	76	.517	76		51	.001	120		77	.245	94
					52	.001	119		78	.273	93
					53	.001	118		79	.302	92
					54	.002	117		80	.333	91
					55	.003	116		81	.365	90
					56	.004	115		82	.398	89
					57	.005	114		83	.432	88
					58	.007	113		84	.466	87
					59	.009	112		85	.500	86
					60	.012	111				
					61	.016	110				
					62	.020	109				
					63	.025	108				
					64	.031	107				
					65	.039	106				
					66	.047	105				
					67	.057	104				
					68	.068	103				
					69	.081	102				
					70	.095	101				

(续附表 X)

n	左 T_x	P	右 T_x
	$m=9$		
10	45	.000	135
	46	.000	134
	47	.000	133
	48	.000	132
	49	.000	131
	50	.000	130
	51	.000	129
	52	.000	128
	53	.001	127
	54	.001	126
	55	.001	125
	56	.002	124
	57	.003	123
	58	.004	122
	59	.005	121
	60	.007	120
	61	.009	119
	62	.011	118
	63	.014	117
	64	.017	116
	65	.022	115
	66	.027	114
	67	.033	113
	68	.039	112
	69	.047	111
	70	.056	110
	71	.067	109
	72	.078	108
	73	.091	107
	74	.106	106
	75	.121	105
	76	.139	104
	77	.158	103

n	左 T_x	P	右 T_x
	$m=9$		
10	78	.178	102
	79	.200	101
	80	.223	100
	81	.248	99
	82	.274	98
	83	.302	97
	84	.330	96
	85	.360	95
	86	.390	94
	87	.421	93
	88	.452	92
	89	.484	91
	90	.516	90
	$m=10$		
10	55	.000	155
	56	.000	154
	57	.000	153
	58	.000	152
	59	.000	151
	60	.000	150
	61	.000	149
	62	.000	148
	63	.000	147
	64	.001	146
	65	.001	145
	66	.001	144
	67	.001	143
	68	.002	142
	69	.003	141
	70	.003	140
	71	.004	139
	72	.006	138

n	左 T_x	P	右 T_x
	$m=10$		
10	73	.007	137
	74	.009	136
	75	.012	135
	76	.014	134
	77	.018	133
	78	.022	132
	79	.026	131
	80	.032	130
	81	.038	129
	82	.045	128
	83	.053	127
	84	.062	126
	85	.072	125
	86	.083	124
	87	.095	123
	88	.109	122
	89	.124	121
	90	.140	120
	91	.157	119
	92	.176	118
	93	.197	117
	94	.218	116
	95	.241	115
	96	.264	114
	97	.289	113
	98	.315	112
	99	.342	111
	100	.370	110
	101	.398	109
	102	.427	108
	103	.456	107
	104	.485	106
	105	.515	105

如果 m 或 n 大于 10，则按下式计算，在附表Ⅳ中查找 P：

$$Z_{x,L}=\frac{T_x+0.5-m(N+1)/2}{\sqrt{mn(N+1)/12}} \qquad Z_{x,R}=\frac{T_x-0.5-m(N+1)/2}{\sqrt{mn(N+1)/12}}$$

预定的	近似的
T_x 的左尾概率	Z_x 的左尾概率
T_x 的右尾概率	$Z_{x,R}$的右尾概率

附表XI：两 样本 *K-S* 检验统计量

m	n	mnD	p	m	n	mnD	p	m	n	mnD	p
2	2	4	.333	3	6	18	.024	4	5	20	.016
2	3	6	.200			15	.095			16	.079
2	4	8	.133			12	.333			15	.143
2	5	10	.095	3	7	21	.017	4	6	24	.010
		8	.286			18	.067			20	.048
2	6	12	.071			15	.167			18	.095
		10	.214	3	8	24	0.12			16	.181
2	7	14	.056			21	0.48	4	7	28	.006
		12	.167			18	.121			24	.030
2	8	16	.044	3	9	27	.009			21	.067
		14	.133			24	.036			20	.121
2	9	18	.036			21	.091	4	8	32	.004
		16	.109			18	.236			28	.020
2	10	20	.030	3	10	30	.007			24	.085
		18	.091			27	.028			20	.222
		16	.182			24	.070	4	9	36	.003
2	11	22	.026			21	.140			32	.014
		20	.077	3	11	33	.005			28	.042
		18	.154			30	.022			27	.062
2	12	24	.022			27	.055			24	.115
		22	.066			24	.110	4	10	40	.002
		20	.132	3	12	36	.004			36	.010
3	3	9	.100			33	.018			32	.030
3	4	12	.057			30	.044			30	.046
		9	.229			27	.088			28	.084
3	5	15	.036			24	.189			26	.126
		12	.143	4	4	16	.029				
						12	.229				

（续附表Ⅺ）

m	n	mnD	p
4	11	44	.001
		40	.007
		36	.022
		33	.035
		32	.063
		29	.098
		28	.144
4	12	48	.001
		44	.005
		40	.016
		36	.048
		32	.112
5	5	25	.008
		20	.079
		15	.357
5	6	30	.004
		25	.026
		24	.048
		20	.108
5	7	35	.003
		30	.015
		28	.030
		25	.066
		23	.116
5	8	40	.002
		35	.009
		32	0.20
		30	.042
		27	0.79
		25	.126
5	9	45	.001
		40	.006
		36	.014
		35	.028
		31	.056
		30	.086
		27	.119

m	n	mnD	p
5	10	50	.001
		45	.004
		40	.019
		35	.061
		30	.166
5	11	55	.000
		50	.003
		45	.010
		44	.014
		40	.029
		39	.044
		35	.074
		34	.106
6	6	36	.002
		30	.026
		24	.143
6	7	42	.001
		36	.008
		35	.015
		30	.038
		29	.068
		28	.091
		24	.147
6	8	48	.001
		42	.005
		40	.009
		36	.023
		34	.043
		32	.061
		30	.093
		28	.139
6	9	54	.000
		48	.003
		45	.006
		42	.014
		39	.028
		36	.061
		33	.095
		30	.176

m	n	mnD	p
6	10	60	.000
		54	.002
		50	.004
		48	.009
		44	.019
		42	.031
		40	.042
		38	.066
		36	.092
		34	.125
7	7	49	.001
		42	.008
		35	.053
		28	.212
7	8	56	.000
		49	.002
		48	.005
		42	.013
		41	.024
		40	.033
		35	.056
		34	.087
		33	.118
7	9	63	.000
		56	.001
		54	.003
		49	.008
		47	.015
		45	.021
		42	.034
		40	.055
		38	.079
		36	.098
		35	.127
8	8	64	.000
		56	.002
		48	.019
		40	.087
		32	.283

(续附表Ⅺ)

$m=n$	mnD 的右尾概率（双侧统计量） .200	.100	.050	.020	.010
9	45	54	54	63	63
10	50	60	70	70	80
11	66	66	77	88	88
12	72	72	84	96	96
13	78	91	91	104	117
14	84	98	112	112	126
15	90	105	120	135	135
16	112	112	128	144	160
17	119	135	136	153	170
18	126	144	162	180	180
19	133	152	171	190	190
20	140	160	180	200	220
	.100	.050	.025	.010	.005
	mnD_+ 或 mnD_- 的近似右尾概率（单侧统计量）				

D 的右尾概率 .200	.100	.050	.020	.010
$1.07\sqrt{N/mn}$	$1.22\sqrt{N/mn}$	$1.36\sqrt{N/mn}$	$1.52\sqrt{N/mn}$	$1.63\sqrt{N/mn}$
.100	.050	.025	.010	.005
D_+ 或 D_- 的右尾概率				

附表 XII：K-W 检验统计量

观测值 H 的相伴概率表*

样本容量			H	p	样本容量			H	p
n_1	n_2	n_3			n_1	n_2	n_3		
2	1	1	2.7000	.500	4	3	2	6.4444	.008
2	2	1	3.6000	.200				6.3000	.011
2	2	2	4.5714	.067				5.4444	.046
			3.7143	.200				5.4000	.051
3	1	1	3.2000	.300				4.5111	.098
3	2	1	4.2857	.100				4.4444	.102
			3.8571	.133	4	3	3	6.7455	.010
3	2	2	5.3572	.029				6.7091	.013
			4.7143	.048				5.7909	.046
			4.5000	.067				5.7273	.050
			4.4643	.105				4.7091	.092
2	3	1	5.1429	.043				4.7000	.101
			4.5714	.100	4	4	1	6.6667	.010
			4.0000	.129				6.1667	.022
3	3	2	6.2500	011				4.9667	.048
			5.3611	.032				4.8667	.054
			5.1389	.061				4.1667	.082
			4.5556	.100				4.0667	.102
			4.2500	.121	4	4	2	7.0364	.006
3	3	3	7.2000	.004				6.8727	.011
			6.4889	.011				5.4545	.046
			5.6889	.029				5.2364	.052
			5.6000	.050				4.5545	.098
			5.0667	.086				4.4455	.103
			4.6222	.100	4	4	3	7.1439	.010
4	1	1	3.5714	.200				7.1364	.011
4	2	1	4.8214	.057				5.5985	.049
			4.5000	.076				5.5758	.051
			4.0179	.114				4.5455	.099
4	2	2	6.0000	.014				4.4773	.102
			5.3333	.033	4	4	4	7.6538	.008
			5.1250	.052				7.5385	.011
			4.4583	.100				5.6923	.049
			4.1667	.105				5.6538	.054
4	3	1	5.8333	.021				4.6539	.097
			5.2083	.050				4.5001	104
			5.0000	.057	5	1	1	3.8571	143
			4.0556	093	5	2	1	5.2500	.036
			3.8889	.129				5.0000	.048

* 本表摘自（并节略）Kruskal, W. H., and Wallis, W. A. 1952. Use of ranks in one-criterion varance analysis. *J. Amer. Statist. Ass.*, 47. 614-617.（本表已将原著者在 J. Amer. Statist. Ass., 48, 910 上所作的勘误包括在内。）

（续附表 X Ⅲ）

n	R	P	n	R	P	n	R	P	n	R	P
9	.383	.156	10	.964	.000	10	.636	.027	10	.309	.193
	.367	.168		.952	.000		.624	.030		.297	.203
	.350	.179		.939	.000		612	.033		.285	.214
	.333	.193		.927	.000		.600	.037		.273	.224
	.317	.205		.915	.000		.588	.040		.261	.235
	.300	.218		.903	.000		.576	.044		.248	.246
	.283	.231		.891	.001		.564	.048		.236	.257
	.267	.247		879	.001		.552	.052		.224	.268
	.250	.260		.867	.001		.539	.057		.212	.280
	.233	.276		.855	.001		.527	.062		.200	.292
	.217	.290		.842	.002		.515	.067		.188	.304
	.200	.307		.830	.002		.503	.072		.176	.316
	.183	.322		.818	.003		.491	.077		.164	.328
	.167	.339		.806	.004		.479	.083		.152	.341
	.150	.354		.794	.004		.467	.089		.139	.354
	.133	.372		.782	.005		.455	.096		.127	.367
	.117	388		.770	.007		.442	.102		.115	.379
	.100	.405		.758	.008		.430	.109		.103	.393
	.083	.422		.745	.009		.418	.116		.091	.406
	.067	.440		.733	.010		.406	.124		.079	.419
	.050	.456		.721	.012		.394	.132		.067	.433
	.033	.474		.709	.013		.382	.139		.055	.446
	017	.491		.697	.015		.370	.148		.042	.459
	.000	.509		.685	.017		.358	.156		.030	.473
10	1.000	.000		.673	.019		.345	.165		.018	.486
	.988	.000		.661	.022		333	.174		.006	.500
	.976	.000		.648	.025		.321	.184			

(续附表 X Ⅱ)

n	单侧检验 R ($-R$) 右尾 (左尾) 概率					
	.100	.050	025	.010	.005	.001
11	.427	.536	.618	.709	.764	.855
12	.406	.503	.587	.678	.734	.825
13	.385	.484	.560	.648	.703	.797
14	.367	.464	.538	.626	.679	.771
15	.354	.446	.521	.604	.657	.750
16	.341	.429	.503	.585	.635	.729
17	.329	.414	.488	.566	.618	.711
18	.317	.401	.474	.550	.600	.692
19	.309	.391	.460	.535	.584	.675
20	.299	.380	.447	.522	.570	.660
21	.292	.370	.436	.509	.556	.647
22	.284	.361	.425	.497	.544	.633
23	.278	.353	.416	.486	.532	.620
24	.275	.344	.407	.476	.521	.608
25	.265	.337	.398	.466	.511	.597
26	.260	.331	.390	.457	.501	.586
27	.255	.324	.383	.449	.492	.576
28	.250	.318	.376	.441	.483	.567
29	.245	.312	.369	.433	.475	.557
30	.241	.307	.363	.426	.467	.548
	.200	.100	.050	.020	.010	.002
	双侧检验 $\|R\|$ 的概率.					

如果 $n>30$，则按 $Z=R\sqrt{n-1}$ 在附表 Ⅳ 中查找概率。

附表 XIV：*kendall* τ 统计量

n	T	p	n	T	p	n	T	p	n	T	p
3	1.000	.167		.714	.015	9	1.000	.000	10	1.000	.000
	.333	.500		.619	.035		.944	.000		.956	.000
4	1.000	.042		.524	.068		.889	.000		.911	.000
	.667	.167		.429	.119		.833	.000		.867	.000
	.333	.375		.333	.191		.778	.001		.822	.000
	.000	.625		.238	.281		.722	.003		.778	.000
5	1.000	.008		.143	.386		.667	.006		.733	.001
	.800	.042		.048	.500		.611	.012		.689	.002
	.600	.117	8	1.000	.000		.556	.022		.644	.005
	.400	.242		.929	.000		.500	.038		.600	.008
	.200	.408		.857	.001		.444	.060		.556	.014
	.000	.592		.786	.003		.389	.090		.511	.023
6	1.000	.001		.714	.007		.333	.130		.467	.036
	.867	.008		.643	.016		.278	.179		.422	.054
	.733	.028		.571	.031		.222	.238		.378	.078
	.600	.068		.500	.054		.167	.306		.333	.108
	.467	.136		.429	.089		.111	.381		.289	.146
	.333	.235		.357	.138		.056	.460		.244	.190
	.200	.360		.286	.199		.000	.540		.200	.242
	.067	.500		.214	.274					.156	.300
7	1.000	.000		.143	.360					.111	.364
	.905	.001		.071	.452					.067	.431
	.810	.005		.000	.548					.022	.500

（续附表 XIV）

n	单侧检验 T (-T) 左尾（左尾）概率				
	.100	.050	.025	.010	.005
11	.345	.418	.491	.564	.600
12	.303	.394	.455	.545	.576
13	.308	.359	.436	.513	.564
14	.275	.363	.407	.473	.516
15	.276	.333	.390	.467	.505
16	.250	.317	.383	.433	.483
17	.250	.309	.368	.426	.471
18	.242	.294	.346	.412	.451
19	.228	.287	.333	.392	.439
20	.221	.274	.326	.379	.421
21	.210	.267	.314	.371	.410
22	.203	.264	.307	.359	.394
23	.202	.257	.396	.352	.391
24	.196	.246	.290	.341	.377
25	.193	.240	.287	.333	.367
26	.188	.237	.280	.329	.360
27	.179	.231	.271	.322	.356
28	.180	.228	.265	.312	.344
29	.172	.222	.261	.310	.340
30	.172	.218	.255	.301	.333
	.200	.100	.050	.020	.010
	双侧检验 $\|T\|$ 的概率				

如果 $n>30$，则按 $Z=3T\sqrt{n(n-1)}/\sqrt{2(2n+5)}$，在附表Ⅳ中查找相应的概率。

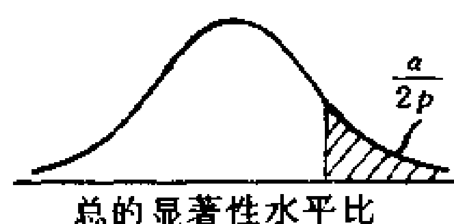
(续附表XV)

若 n 和 k 的值在表中范围之外，则按下式在附表 I 中得右尾概率：

$$Q=\frac{12S}{kn(n+1)}$$

预定的	近似的
S 或 W 的 左尾概率	$df=n-1$ 的 Q 的左尾概率

附表 XVI：多重比较的临界值 Z



总的显著性水平 α

P	α .30	.25	.20	.15	.10	.05
1	1.036	1.150	1.282	1.440	1.645	1.960
2	1.440	1.534	1.645	1.780	1.960	2.241
3	1.645	1.732	1.834	1.960	2.128	2.394
4	1.780	1.863	1.960	2.080	2.241	2.498
5	1.881	1.960	2.054	2.170	2.326	2.576
6	1.960	2.037	2.128	2.241	2.394	2.638
7	2.026	2.100	2.189	2.300	2.450	2.690
8	2.080	2.154	2.241	2.350	2.498	2.734
9	2.128	2.200	2.287	2.394	2.539	2.773
10	2.170	2.241	2.326	2.432	2.576	2.807
11	2.208	2.278	2.362	2.467	2.608	2.838
12	2.241	2.301	2.394	2.498	2.638	2.866
15	2.326	2.394	2.475	2.576	2.713	2.935
21	2.450	2.515	2.593	2.690	2.823	3.038
28	2.552	2.615	2.690	2.785	2.913	3.125

附表 XVII：F 分布表

$$P\{F(n_1,n_2) > F\alpha(n_1,n_2)\} = \alpha$$

$\alpha=0.10$

n_2 \ n_1	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	39.86	49.50	53.59	55.83	57.24	58.20	58.91	59.44	59.86	60.19	60.71	61.22	61.74	62.00	62.26	62.53	62.79	63.06	63.33
2	8.53	9.00	9.16	9.24	9.29	9.33	9.35	9.37	9.38	9.39	9.41	9.42	9.44	9.45	9.46	9.47	9.47	9.48	9.49
3	5.54	5.46	5.39	5.34	5.31	5.28	5.27	5.25	5.24	5.23	5.22	5.20	5.18	5.18	5.17	5.16	5.15	5.14	5.13
4	4.54	4.32	4.19	4.11	4.05	4.01	3.98	3.95	3.94	3.92	3.90	3.87	3.84	3.83	3.82	3.80	3.79	3.78	3.76
5	4.06	3.78	3.62	3.52	3.45	3.40	3.37	3.34	3.32	3.30	3.27	3.24	3.21	3.19	3.17	3.16	3.14	3.12	3.10
6	3.78	3.46	3.29	3.18	3.11	3.05	3.01	2.98	2.96	2.94	2.90	2.87	2.84	2.82	2.80	2.78	2.76	2.74	2.72
7	3.59	3.26	3.07	2.96	2.88	2.83	2.78	2.75	2.72	2.70	2.67	2.63	2.59	2.58	2.56	2.54	2.51	2.49	2.47
8	3.46	3.11	2.92	2.81	2.73	2.67	2.62	2.59	2.56	2.54	2.50	2.46	2.42	2.40	2.38	2.36	2.34	2.32	2.20
9	3.36	3.01	2.81	2.69	2.61	2.55	2.51	2.47	2.44	2.42	2.38	2.34	2.30	2.28	2.25	2.23	2.21	2.18	2.16
10	3.29	2.92	2.73	2.61	2.52	2.46	2.41	2.38	2.35	2.32	2.28	2.24	2.20	2.18	2.16	2.13	2.11	2.08	2.06
11	3.23	2.86	2.66	2.54	2.45	2.39	2.34	2.30	2.27	2.25	2.21	2.17	2.12	2.12	2.08	2.05	2.03	2.00	1.97
12	3.18	2.81	2.61	2.48	2.39	2.33	2.28	2.24	2.21	2.19	2.15	2.10	2.04	5.06	2.01	1.99	1.96	1.93	1.90
13	3.14	2.76	2.56	2.43	2.35	2.28	2.23	2.20	2.16	2.14	2.10	2.05	2.01	1.98	1.96	1.93	1.90	1.88	1.85
14	3.10	2.73	2.52	2.39	2.31	2.24	2.19	2.15	2.12	2.10	2.05	2.01	1.96	1.94	1.91	1.89	1.86	1.83	1.80
15	3.07	2.70	2.49	2.36	2.27	2.21	2.16	2.12	2.09	2.06	2.07	1.97	1.92	1.90	1.87	1.85	1.82	1.79	1.76
16	3.05	2.67	2.46	2.33	2.24	2.18	2.13	2.09	2.06	2.03	1.99	1.94	1.89	1.87	1.84	1.81	1.78	1.75	1.72
17	3.03	2.64	2.44	2.31	2.22	2.16	2.10	2.06	2.03	2.00	1.96	1.91	1.84	1.84	1.81	1.78	1.75	1.73	1.69
18	3.01	2.62	2.42	2.29	2.20	2.13	2.08	2.04	2.00	1.98	1.93	1.89	1.81	1.81	1.78	1.75	1.72	1.69	1.66
19	2.99	2.61	2.40	2.27	2.18	2.11	2.06	2.02	1.98	1.96	1.91	1.86	1.81	1.79	1.76	1.73	1.70	1.67	1.63

(续附表 XVII)

$\alpha=0.10$

n_2 \ n_1	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
20	2.97	2.59	2.38	2.25	2.16	2.09	2.04	2.00	1.96	1.94	1.89	1.84	1.79	1.77	1.74	1.71	1.63	1.64	1.61
21	2.96	2.57	2.36	2.23	2.14	2.08	2.02	1.98	1.95	1.92	1.87	1.83	1.78	1.75	1.72	1.69	1.66	1.62	1.59
22	2.95	2.56	2.35	2.22	2.13	2.06	2.01	1.97	1.93	1.90	1.86	1.81	1.76	1.73	1.70	1.67	1.64	1.60	1.57
23	2.94	2.55	2.34	2.21	2.11	2.05	1.99	1.95	1.92	1.89	1.84	1.80	1.74	1.72	1.69	1.66	1.62	1.59	1.55
24	2.93	2.54	2.33	2.19	2.10	2.04	1.98	1.94	1.91	1.88	1.83	1.78	1.73	1.70	1.67	1.64	1.61	1.57	1.53
25	2.92	2.53	2.32	2.18	2.09	2.02	1.97	1.93	1.89	1.87	1.82	1.77	1.72	1.69	1.66	1.63	1.59	1.56	1.52
26	2.91	2.52	2.31	2.17	2.08	2.01	1.96	1.92	1.88	1.86	1.81	1.76	1.71	1.68	1.65	1.61	1.58	1.54	1.50
27	2.90	2.51	2.30	2.17	2.07	2.00	1.95	1.91	1.87	1.85	1.80	1.75	1.70	1.67	1.64	1.60	1.57	1.53	1.49
28	2.89	2.50	2.29	2.16	2.06	2.00	1.94	1.90	1.87	1.84	1.79	1.74	1.69	1.66	1.63	1.59	1.56	1.52	1.48
29	2.89	2.50	2.28	2.15	2.06	1.99	1.93	1.89	1.86	1.83	1.78	1.73	1.68	1.65	1.62	1.58	1.55	1.51	1.47
30	2.88	2.49	2.28	2.14	2.05	1.98	1.93	1.88	1.85	1.82	1.77	1.72	1.67	1.64	1.61	1.57	1.54	1.50	1.46
40	2.84	2.44	2.23	2.09	2.00	1.93	1.87	1.83	1.79	1.76	1.71	1.66	1.61	1.57	1.54	1.51	1.47	1.42	1.38
60	2.79	2.39	2.18	2.04	1.95	1.87	1.82	1.77	1.74	1.71	1.66	1.60	1.54	1.51	1.48	1.44	1.40	1.35	1.29
120	2.75	2.35	2.13	1.99	1.90	1.82	1.77	1.72	1.68	1.65	1.60	1.55	1.48	1.45	1.41	1.37	1.32	1.26	1.19
∞	2.71	2.30	2.08	1.94	1.85	1.77	1.72	1.67	1.63	1.60	1.55	1.49	1.42	1.38	1.34	1.30	1.24	1.17	1.00

(续附表 XVII)

$\alpha=0.05$

n_2 \ n_1	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	161.4	199.5	215.7	224.6	230.2	234.0	236.8	238.9	240.5	241.9	243.9	245.9	248.0	249.1	250.1	251.1	252.2	253.3	254.3
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.41	19.43	19.45	19.45	19.46	19.47	19.48	19.49	19.50
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.70	8.66	8.64	8.62	8.69	8.57	8.55	8.53
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.86	5.80	5.77	5.75	5.72	5.69	5.66	5.63
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.53	4.50	4.46	4.43	4.40	4.36
6	6.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.94	3.87	3.84	3.81	3.77	3.74	3.70	3.67
7	5.59	4.74	4.35	4.12	3.97	3.87	3.70	3.73	3.68	3.64	3.57	3.51	3.44	3.41	3.38	3.34	3.30	3.27	3.23
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.22	3.15	3.12	3.08	3.04	3.01	2.97	2.93
9	5.12	4.20	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01	2.94	2.90	2.86	2.83	2.79	2.75	2.71
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.85	2.77	2.74	2.70	2.66	2.62	2.58	2.54
11	4.84	3.98	3.59	3.36	3.20	3.09	3.10	2.95	2.90	2.85	2.79	2.72	2.65	2.61	2.57	2.53	2.49	2.45	2.40
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.62	2.54	2.51	2.47	2.43	2.38	2.34	2.30
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.60	2.53	2.46	2.42	2.38	2.34	2.30	2.25	2.21
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53	2.46	2.39	2.35	2.31	2.27	2.22	2.18	2.13
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.40	2.33	2.29	2.25	2.20	2.16	2.11	2.07
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.35	2.28	2.24	2.19	2.15	2.11	2.06	2.01
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.38	2.31	2.23	2.19	2.15	2.10	2.06	2.01	1.96
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.27	2.19	2.15	2.11	2.06	2.02	1.97	1.92
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.93	1.88

（续附表 XVII）

$\alpha=0.05$

n_2 \ n_1	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.20	2.12	2.08	2.04	1.99	1.95	1.90	1.84
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.25	2.18	2.10	2.05	2.01	1.96	1.92	1.87	1.81
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.23	2.15	2.07	2.03	1.98	1.94	1.89	1.84	1.78
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2.20	2.13	2.05	2.01	1.96	1.91	1.86	1.81	1.76
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.18	2.11	2.03	1.98	1.94	1.89	1.84	1.79	1.73
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.16	2.09	2.01	1.96	1.92	1.87	1.82	1.77	1.71
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.15	2.07	1.99	1.95	1.90	1.85	1.80	1.75	1.69
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.20	2.13	2.06	1.97	1.93	1.88	1.84	1.79	1.73	1.67
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19	2.12	2.04	1.96	1.91	1.87	1.82	1.77	1.71	1.65
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22	2.18	2.10	2.03	1.94	1.90	1.85	1.81	1.75	1.70	1.64
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09	2.01	1.93	1.89	1.84	1.79	1.74	1.68	1.62
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.00	1.92	1.84	1.79	1.74	1.69	1.64	1.58	1.51
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.92	1.84	1.75	1.70	1.65	1.59	1.53	1.47	1.39
120	3.92	3.07	2.68	2.45	2.29	2.17	2.09	2.02	1.96	1.91	1.83	1.75	1.66	1.61	1.55	1.50	1.43	1.35	1.25
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.75	1.67	1.57	1.52	1.46	1.39	1.32	1.21	1.00

（续附表 XVII）

$\alpha=0.025$

n_2 \ n_1	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	647.8	799.5	864.2	899.6	921.8	937.1	489.2	956.7	963.3	968.6	976.7	984.9	993.1	997.2	1001	1006	1010	1014	1018
2	38.51	39.00	39.17	39.25	39.30	39.33	39.36	39.37	39.39	39.40	39.41	39.43	39.45	39.46	39.46	39.47	39.48	39.49	39.50
3	17.44	16.04	15.44	15.10	14.88	14.73	14.62	14.54	14.47	14.42	14.34	14.25	14.17	14.12	14.08	14.04	13.99	13.95	13.90
4	12.22	10.65	9.98	9.60	9.36	9.20	9.07	8.98	8.90	8.84	8.75	8.66	8.56	8.51	8.46	8.41	8.36	8.31	8.26
5	10.01	8.43	7.76	7.36	7.15	6.98	6.85	6.76	6.68	6.62	6.52	6.43	6.33	6.28	6.23	6.18	6.12	6.07	6.02
6	8.81	7.26	6.60	6.23	5.99	5.82	5.70	5.60	5.52	5.46	5.37	5.27	5.17	5.12	5.07	5.01	4.96	4.90	4.85
7	8.07	6.54	5.89	5.52	5.29	5.12	4.99	4.90	4.82	4.76	4.67	4.57	4.47	4.42	4.36	4.31	4.25	4.20	4.14
8	7.57	6.06	5.42	5.05	4.82	4.65	4.53	4.43	4.36	4.30	4.20	4.10	4.00	3.95	3.89	3.84	3.78	3.73	3.67
9	7.21	5.71	5.08	4.72	4.48	4.32	4.20	4.10	4.03	3.96	3.87	3.77	3.67	3.61	3.56	3.51	3.45	3.39	3.33
10	6.94	5.46	4.83	4.47	4.24	4.07	3.95	3.85	3.78	3.72	3.62	3.52	3.42	3.37	3.31	3.26	3.20	3.14	3.08
11	6.72	5.26	4.63	4.28	4.04	3.88	3.76	3.66	3.59	3.53	3.43	3.38	3.23	3.17	3.12	3.06	3.00	2.94	2.88
12	6.55	5.10	4.47	4.12	3.89	3.73	3.61	3.51	3.44	3.37	3.28	3.18	3.07	3.02	2.96	2.91	2.85	2.79	2.72
13	6.41	4.97	4.35	4.00	3.77	3.60	3.48	3.39	3.31	3.25	3.15	3.05	2.95	2.89	2.84	2.78	2.72	2.66	2.60
14	6.30	4.86	4.24	3.89	3.66	3.50	3.38	3.29	3.21	3.15	3.05	2.95	2.84	2.79	2.73	2.67	2.61	2.55	2.49
15	6.20	4.77	4.15	3.80	3.58	3.41	3.29	3.20	3.12	3.06	2.96	2.86	2.76	2.70	2.64	2.59	2.52	2.46	2.40
16	6.12	4.69	4.08	3.73	3.50	3.34	3.22	3.12	3.05	2.99	2.89	2.79	2.68	2.63	2.57	2.51	2.45	2.38	2.32
17	6.04	4.62	4.01	3.66	3.44	3.28	3.16	3.06	2.98	2.92	2.82	2.72	2.62	2.56	2.50	2.44	2.38	2.32	2.25
18	5.98	4.56	3.95	3.61	3.38	3.22	3.10	3.01	2.93	2.87	2.77	2.67	2.56	2.50	2.44	2.38	2.32	2.26	2.19
19	5.92	4.51	3.90	3.56	3.33	3.17	3.05	2.96	2.88	2.82	2.72	2.62	2.51	2.45	2.39	2.33	2.27	2.20	2.13
20	5.87	4.46	3.86	3.51	3.29	3.13	3.01	2.91	2.84	2.77	2.68	2.57	2.46	2.41	2.35	2.29	2.22	2.16	2.09
21	5.83	4.42	3.82	3.48	3.25	3.09	2.97	2.87	2.80	2.73	2.64	2.56	2.42	2.37	2.31	2.25	2.18	2.11	2.04
22	5.79	4.38	3.78	3.44	3.22	3.05	2.93	2.84	2.76	2.70	2.60	2.50	2.39	2.33	2.27	2.21	2.14	2.08	2.00
23	5.75	4.35	3.75	3.41	3.18	3.02	2.90	2.81	2.73	2.67	2.57	2.47	2.36	2.30	2.24	2.18	2.11	2.04	1.97
24	5.72	4.32	3.72	3.38	3.15	2.99	2.87	2.78	2.70	2.64	2.54	2.44	2.33	2.27	2.21	2.15	2.08	2.01	1.91

(续附表 XVII)　　　　$a=0.01$

n_2 \ n_1	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
10	0.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85	4.71	4.50	4.41	4.33	4.25	4.17	4.08	4.00	3.91
11	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63	4.54	4.40	4.25	4.10	4.02	3.94	3.86	4.78	3.69	3.60
12	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39	4.30	4.16	4.01	3.86	3.78	3.70	3.62	3.54	3.45	3.36
13	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19	3.10	3.96	3.82	3.66	3.59	3.51	3.43	3.34	3.25	3.17
14	8.86	6.51	5.56	5.04	4.69	4.46	4.28	4.14	4.03	3.94	3.80	3.66	3.51	3.43	3.35	3.27	3.18	3.09	3.00
15	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89	3.80	3.67	3.52	3.37	3.29	3.21	3.13	3.05	2.96	2.87
16	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78	3.69	3.55	3.41	3.26	3.18	3.10	3.02	2.93	2.84	2.75
17	8.40	6.11	5.18	4.67	4.34	4.10	3.93	3.79	3.68	3.59	3.46	3.31	3.16	3.08	3.00	2.92	2.83	2.75	2.65
18	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60	3.51	3.37	3.23	3.08	3.00	2.92	2.84	2.75	2.66	2.57
19	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52	3.43	3.30	3.15	3.00	7.92	2.84	2.76	2.67	2.58	2.49
20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46	3.37	3.23	3.09	2.94	2.86	2.78	2.69	2.61	2.52	2.42
21	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40	3.31	3.17	3.03	2.88	2.80	2.72	2.64	2.55	2.46	2.36
22	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35	3.26	3.12	2.98	2.83	2.75	2.67	2.58	2.50	2.40	2.31
23	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30	3.21	3.07	2.93	2.78	2.70	2.62	2.54	2.45	2.35	2.26
24	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26	3.17	3.03	2.89	2.74	2.66	2.58	2.49	2.40	2.31	2.21
25	7.77	5.57	4.68	4.18	3.85	3.63	3.46	3.32	3.22	3.13	2.99	2.85	2.70	2.62	2.54	2.45	2.36	2.27	2.17
26	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.18	3.09	2.96	2.81	2.66	2.58	2.50	2.42	2.33	2.23	2.13
27	7,64	5.49	4.60	4.11	3.78	3.56	3.39	3.26	3.15	3.06	2.93	2.78	2.63	2.55	2.47	2.38	2.29	2.20	2.10
28	7.68	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.12	3.03	2.90	2.75	2.00	2.52	2.44	2.35	2.26	2.17	2.06
29	7.60	5.42	4.54	4.04	3.73	3.50	3.33	3.20	3.09	3.00	2.87	2.73	2.57	2.49	2.41	2.33	2.23	2.14	2.03
30	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07	2.98	2.84	2.70	2.55	2.47	2.39	2.30	2.21	2.11	2.01
40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89	2.80	2.66	2.52	2.37	2.29	2.20	2.11	2.02	1.92	1.80
60	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63	2.50	2.35	2.20	2.12	2.03	1.94	1.84	1.73	1.60
120	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56	2.47	2.34	2.19	2.00	1.95	1.86	1.76	1.66	1.53	1.38
∞	6.63	4.01	3.78	3.32	3.02	2.80	2.64	2.51	2.41	2.32	2.18	2.04	1.88	1.79	1.70	1.59	1.47	1.32	1.00

(续附表 XVII)

$\alpha=0.005$

n_2 \ n_1	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	16211	20000	21615	22500	23056	23437	23715	23925	24091	24224	24426	24630	24836	24940	25044	25148	25253	25339	25465
2	198.5	199.0	199.2	199.2	199.3	199.3	199.4	199.4	199.4	199.4	199.4	199.4	199.4	199.5	199.5	199.5	199.5	199.5	199.5
3	55.55	49.80	47.47	46.19	45.39	44.84	44.43	44.13	43.88	43.69	43.39	43.08	42.78	42.62	42.47	42.31	42.15	41.99	41.83
4	31.33	26.28	24.26	23.15	22.46	21.79	21.62	21.35	21.14	20.97	20.70	20.44	20.17	20.03	19.89	19.75	19.61	19.47	19.32
5	22.78	18.31	16.53	15.56	14.94	14.51	14.20	13.96	13.77	13.62	13.38	13.15	12.90	12.78	12.66	12.53	12.40	12.27	12.14
6	18.63	14.54	12.92	12.03	11.46	11.07	10.79	10.57	10.39	10.25	10.03	9.81	9.59	9.47	9.36	9.24	9.12	9.00	8.38
7	16.24	12.40	10.88	10.05	9.52	9.16	8.89	8.68	8.51	8.38	8.81	7.97	7.75	7.65	7.53	7.42	7.31	7.19	7.08
8	14.69	11.04	9.60	8.81	8.30	7.95	7.69	7.50	7.34	7.21	7.01	6.81	6.61	6.50	6.40	6.29	6.18	6.06	5.95
9	13.61	10.11	8.72	7.96	7.47	7.13	6.88	6.69	6.54	6.42	6.23	6.03	5.83	5.73	5.62	5.52	5.41	5.30	7.19
10	12.83	9.43	8.08	7.34	6.87	6.54	6.30	6.12	5.97	5.85	5.66	5.47	5.27	5.17	5.07	4.97	4.80	4.75	4.34
11	12.23	8.91	7.60	6.88	6.42	6.10	5.86	5.68	5.54	5.42	5.24	5.05	4.86	4.76	4.65	4.55	4.44	4.34	4.23
12	11.75	8.51	7.23	6.52	6.07	5.76	5.52	5.35	5.20	5.09	4.91	4.72	4.53	4.43	4.33	4.23	4.12	4.01	3.90
13	11.37	8.19	6.93	6.23	5.79	5.48	5.25	5.08	4.94	4.82	4.64	4.46	4.27	4.17	4.07	3.97	3.87	3.70	3.65
14	11.06	7.92	6.68	6.00	5.56	5.26	5.03	4.80	4.72	4.60	4.43	4.25	4.06	3.90	3.86	3.76	3.66	3.55	3.44
15	10.80	7.70	6.48	5.80	5.37	5.07	4.85	4.67	4.54	4.42	4.25	4.07	3.88	3.79	3.69	3.58	3.48	3.37	3.28
16	10.58	7.51	6.30	5.64	5.21	4.91	4.69	4.52	4.38	4.27	4.10	3.92	3.73	3.64	3.54	3.44	3.33	3.22	3.11
17	10.38	7.35	6.16	5.50	5.07	4.78	4.56	4.39	4.25	4.14	3.97	3.79	3.61	3.51	3.41	3.31	3.21	3.10	2.98
18	10.22	7.21	6.03	5.37	4.96	4.66	4.44	4.28	4.14	4.03	3.86	3.68	3.50	3.40	3.30	3.20	3.10	2.99	2.87
19	10.07	7.09	5.92	5.27	4.85	4.56	4.34	4.18	4.04	3.93	3.76	3.59	3.40	3.31	3.21	3.11	3.00	2.89	2.78
20	9.94	6.90	5.82	5.17	4.76	4.47	4.26	4.09	3.96	3.85	3.68	3.50	3.32	3.22	3.12	3.02	2.92	2.81	2.69
21	9.83	6.89	5.73	5.09	4.68	4.30	4.18	4.01	3.88	3.77	3.60	3.48	3.24	3.15	3.05	2.95	2.84	2.73	2.61
22	9.73	6.81	5.65	5.02	4.61	4.32	4.11	3.94	3.81	3.70	3.54	3.36	3.18	3.08	2.98	2.88	2.77	2.66	2.55
23	9.63	6.73	5.58	4.95	4.54	4.26	4.05	3.88	3.75	3.64	3.47	3.30	3.12	3.02	2.92	2.82	2.71	2.60	2.48
24	9.55	6.66	5.52	4.89	4.49	4.20	3.99	3.83	3.69	3.59	3.42	3.25	3.06	2.97	2.87	2.77	2.66	2.55	2.43

(续附表 XVII)

$\alpha=0.005$

n_2 \ n_1	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
25	9.48	6.60	5.46	4.84	4.43	4.15	3.94	3.78	3.64	3.54	3.37	3.20	3.01	2.92	2.82	2.72	2.61	2.50	2.38
26	9.41	6.54	5.41	4.79	4.38	4.10	3.89	3.73	3.60	3.49	3.33	3.15	2.97	2.87	2.77	2.67	2.56	2.45	2.33
27	9.34	6.49	5.36	4.74	4.34	4.06	3.85	3.69	3.56	3.45	3.28	3.11	2.93	2.83	2.73	2.63	2.52	2.44	2.29
28	9.28	6.44	5.32	4.70	4.30	4.02	3.81	3.65	3.52	3.41	3.25	3.07	2.89	2.79	2.69	2.59	2.48	2.37	2.25
29	9.23	6.40	5.28	4.66	4.26	3.98	3.77	3.61	3.48	3.38	3.21	3.04	2.86	2.76	2.66	2.56	2.45	2.33	2.21
30	9.18	6.35	5.24	4.62	4.23	3.95	3.74	3.58	3.45	3.34	3.18	3.01	2.82	2.73	2.63	2.52	2.42	2.30	2.18
40	8.83	6.07	4.98	4.37	3.99	3.71	3.51	3.35	3.22	3.12	2.95	2.78	2.60	2.50	2.40	2.30	2.18	2.06	1.93
60	8.49	5.79	4.73	4.14	3.76	3.49	3.29	3.13	3.01	2.90	2.74	2.57	2.39	2.29	2.19	2.08	1.96	1.83	1.69
120	8.18	5.54	4.50	3.92	3.55	3.28	3.09	2.93	2.81	2.71	2.54	2.37	2.19	2.09	1.98	1.87	1.75	1.61	1.43
∞	7.88	5.30	4.28	3.72	3.35	3.09	2.90	2.74	2.62	2.52	2.36	2.19	2.00	1.90	1.79	1.67	1.53	1.36	1.00

$\alpha=0.001$

n_2 \ n_1	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	4053	5000	5404	5625	5764	5859	5929	5981	6023	6056	6107	6158	6209	6235	6261	6287	6313	6340	6366
2	998.5	999.0	999.2	999.2	999.3	999.3	999.4	999.4	999.4	999.4	999.4	999.4	999.4	999.5	999.5	999.5	999.5	999.5	999.5
3	167.0	148.5	141.1	137.1	134.6	132.8	131.6	130.6	129.9	129.2	128.3	127.4	126.4	12.59	125.4	125.0	124.5	124.0	123.5
4	74.14	61.25	56.18	53.44	51.71	50.53	49.66	49.00	48.47	48.05	47.41	46.76	46.10	45.77	45.43	45.09	44.75	44.40	44.05
5	47.18	37.12	33.20	31.09	29.75	28.84	28.16	27.64	27.24	26.92	26.42	25.91	25.36	25.14	24.87	24.60	24.33	24.06	23.78
6	35.51	27.00	23.70	21.92	20.81	20.03	19.46	19.03	18.69	18.41	17.99	17.56	17.12	16.89	16.67	16.44	16.21	15.99	15.75
7	29.25	21.69	18.77	17.19	16.21	15.52	15.62	14.63	14.33	14.08	13.71	13.32	12.93	12.73	12.53	12.33	12.12	11.91	11.70
8	25.42	18.49	15.83	14.39	13.49	12.36	12.40	12.04	11.77	11.54	11.19	10.84	10.48	10.30	10.11	9.92	9.73	9.53	9.33
9	22.86	16.39	13.90	12.56	11.71	11.13	10.70	10.37	10.11	9.89	9.57	9.24	8.90	8.72	8.55	8.37	8.19	8.00	7.81

(续附表 XVII)

$\alpha=0.001$

n_2 \ n_1	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
10	21.04	14.91	12.55	11.28	10.48	9.92	9.52	9.20	8.96	8.75	8.45	8.13	7.80	7.64	7.47	7.30	7.12	6.94	6.76
11	19.69	13.81	11.56	10.35	9.58	9.05	8.66	8.35	8.12	7.92	7.63	7.32	7.01	6.85	6.68	6.52	6.35	6.17	6.00
12	18.64	12.97	10.80	8.63	9.89	8.38	8.00	7.71	7.48	7.29	7.00	6.71	6.40	6.25	6.09	5.93	5.76	6.59	5.42
13	17.81	12.31	10.21	9.07	8.35	7.86	7.49	7.21	6.98	6.80	6.52	6.23	5.93	5.70	5.63	5.47	5.30	5.14	4.97
14	17.14	11.78	9.73	8.62	7.92	7.43	7.08	6.80	6.58	6.40	6.13	5.85	5.56	5.41	5.25	5.10	4.94	4.77	4.60
15	16.59	11.34	9.34	8.25	7.57	7.09	6.74	6.47	6.26	6.08	5.81	5.54	5.25	5.10	4.95	4.80	4.64	4.47	4.31
16	16.12	10.97	9.00	7.94	7.27	6.81	6.46	6.19	5.98	5.81	5.55	5.27	4.99	4.85	4.70	4.54	4.39	4.26	4.06
17	16.72	10.66	8.73	7.68	7.02	7.56	6.22	5.96	5.75	5.58	5.32	5.05	4.78	4.63	4.48	4.33	4.18	4.02	3.85
18	15.38	10.39	8.40	7.46	6.81	6.35	6.02	5.76	5.56	5.39	5.13	4.87	4.59	4.45	4.30	4.15	4.00	3.84	3.67
19	15.03	10.16	8.28	7.26	6.62	6.18	5.85	5.59	5.39	5.22	4.97	4.70	4.43	4.29	4.14	3.99	3.84	3.63	3.51
20	14.82	9.95	8.10	7.10	6.46	6.02	5.69	5.44	5.24	5.08	4.82	4.56	4.29	4.15	4.00	3.86	3.70	3.54	3.38
21	14.59	9.77	7.94	6.95	6.32	5.88	5.56	5.31	5.11	4.95	4.70	4.44	4.17	4.03	3.88	3.74	3.58	3.42	3.26
22	14.38	9.61	7.80	6.81	6.19	5.76	5.44	5.19	5.99	4.83	4.58	4.33	4.06	3.92	3.78	3.63	3.48	3.32	3.15
23	14.19	9.47	7.67	6.69	6.08	5.65	5.33	5.09	4.89	4.73	4.48	4.23	3.96	3.82	3.68	8.53	3.38	3.22	3.05
24	14.03	9.34	7.55	6.59	5.98	5.55	5.23	4.99	4.80	4.64	4.39	4.14	3.87	3.74	3.59	3.45	3.29	3.14	2.97
25	13.88	9.22	7.45	6.49	5.88	5.46	5.15	4.91	4.71	4.56	4.31	4.06	3.79	3.66	3.52	3.37	3.22	3.06	2.89
26	13.74	9.12	7.36	6.41	5.80	5.38	5.07	4.83	4.64	4.48	4.24	3.99	3.72	3.59	3.44	3.30	3.15	2.99	2.82
27	13.61	9.02	7.27	6.33	5.73	5.31	5.00	4.76	4.57	4.41	4.17	3.92	3.66	3.52	3.38	3.23	3.08	2.92	2.75
28	13.50	8.93	7.19	6.25	5.66	5.24	4.93	4.69	4.50	4.35	4.11	3.86	3.60	3.46	3.32	3.18	3.02	2.86	2.69
29	13.39	8.85	7.12	6.19	5.59	5.18	4.87	4.64	4.45	4.29	4.05	3.80	3.54	3.41	3.27	3.12	2.97	2.81	2.64
30	13.29	8.77	7.05	6.12	5.53	5.12	4.82	4.58	4.39	4.24	4.00	3.75	3.49	3.36	3.22	3.07	2.92	2.76	2.59
40	12.61	8.25	6.00	5.70	5.13	4.73	4.44	4.21	4.02	3.87	3.64	3.40	3.15	3.01	2.87	2.73	2.57	2.41	2.23
60	11.97	7.76	6.17	5.31	4.76	4.37	4.09	3.87	3.69	3.54	3.31	3.08	2.83	2.69	2.55	2.41	2.25	2.08	1.89
120	11.38	7.32	5.79	4.95	4.42	4.04	3.77	3.55	3.38	3.24	3.02	2.78	2.53	2.40	2.26	2.11	1.95	1.76	1.54
∞	10.83	6.91	5.42	4.62	4.10	3.74	3.47	3.27	3.10	2.96	2.74	2.51	2.27	2.13	1.99	1.84	1.66	1.45	1.00